
AGGA: A Dataset of Academic Guidelines for Generative AIs

Junfeng Jiao^{1*} Saleh Afroogh^{*2} Kevin Chen³ David Atkinson⁴ Amit Dhurandhar⁵

1. Urban Information Lab, The School of Architecture, The University of Texas at Austin, Austin, TX 78712, United States.
jjiao@austin.utexas.edu

2. Urban Information Lab, The School of Architecture, The University of Texas at Austin, Austin, TX 78712, United States.
Saleh.afroogh@utexas.edu

3. Kevin Chen, Department of economics, The University of Texas at Austin, Austin, TX 78712, USA. xc4646@utexas.edu

4. David Atkinson, Allen Institute for AI (AI2), Seattle, USA davida@allenai.org

5. Amit Dhurandhar, IBM Research Yorktown Heights, USA, adhuran@us.ibm.com

* Corresponding author: saleh.afroogh@utexas.edu

AGGA (Academic Guidelines for Generative AIs) is a dataset of 80 academic guidelines for the usage of generative AIs and large language models in academia, selected systematically and collected from official university websites across six continents. Comprising 181,225 words, the dataset supports natural language processing tasks such as language modeling, sentiment and semantic analysis, model synthesis, classification, and topic labeling. It can also serve as a benchmark for ambiguity detection and requirements categorization. This resource aims to facilitate research on AI governance in educational contexts, promoting a deeper understanding of the integration of AI technologies in academia.

For using any of the described data in parts or in full or for quoting this note, please cite the resource as:

Jiao, Junfeng, Saleh Afroogh, Kevin Chen, David Atkinson, and Amit Dhurandhar. 2024. ‘AGGA: A Dataset of Academic Guidelines for Generative AIs’. Harvard Dataverse. <https://doi.org/10.7910/DVN/XZZHAA5>.

This dataset is organized into several sections to facilitate its use: **Section I** contains a table listing the 80 academic guidelines; **Section II** provides academic citations for each guideline; and **Section III** includes the complete texts of all 80 guidelines, accompanied by a detailed table of contents.

I : 80 Academic Guidelines for Generative AIs

	Continent	Country	University	Name of document/website
1	Africa	South Africa	University of Cape Town	Artificial Intelligence for Teaching & Learning [1]
2		Egypt	Cairo University	FCAI Policy and Guidelines for use of Generative AI in Postgraduate Studies and Research [2]
3		South Africa	University of Witwatersrand	ChatGPT & other AI tools for Learning and Teaching [3]
4		South Africa	University of Pretoria	Leveraging Generative Artificial Intelligence for Teaching and Learning Enhancement [4]
5		South Africa	University of Johannesburg	UJ Practice Notes: Generative Artificial Intelligence in Teaching, Learning and Research [5]
6		South Africa	Stellenbosch University	Stellenbosch University Academic Integrity: Responsible Use of AI tools [6]
7		South Africa	University of the Free State	Stepping up with ChatGPT - AI-assisted Technology in Education [7]
8		Egypt	The American University in Cairo	AUC's Statement on the Use of Artificial Intelligence Tools [8]
9		South Africa	North-West University	Implications of AI for teaching and learning in higher education & Guidelines for the Utilization of AI in Teaching and Learning at NWU [9]
10		South Africa	African Observatory on Responsible Artificial Intelligence	Generative AI guidelines at South African universities [10]
11	North America	USA	Harvard	Guidelines for Using ChatGPT and other Generative AI tools at Harvard [11]
12		USA	Stanford	Generative AI Policy Guidance [12]
13		USA	MIT	Getting Started with AI-Enhanced Teaching: A Practical Guide for Instructors [13]
14		USA	Princeton	Generative AI Guidance [14]

15	North America	USA	University of Chicago	Guidance for Syllabus Staements on the Use of AI Tools [15]
16		USA	Columbia	Considerations for AI Tools in the Classroom [16]
17		USA	California Institute of Technology	Guidance on the Use of Generative AI and Large Language Model Tools [17]
18		USA	University of California, Berkeley	Appropriate use of ChatGPT and Similar AI Tools [18]
19		USA	Yale	Guidelines for the Use of Generative AI Tools [19]
20		USA	University of Pennsylvania	Statement on Guidance for the University of Pennsylvania Community on Use of Generative Artificial Intelligence [20]
21		USA	University of California, Los Angeles	Guidance for the use of generative AI [21]
22		USA	Cornell University	Cornell Guidelines for artificial intelligence [22]
23		Canada	University of Toronto	Generative Artificial Intelligence in the classroom [23]
24		Canada	University of British Columbia	Generative AI – Academic Integrity at UBC [24]
25		Canada	McGill University	Principles on Generative AI in Teaching and Learning at McGill [25]
26		Canada	University of Alberta	AI-Squared – Artificial Intelligence and Academic Integrity [26]
27		Canada	University of Waterloo	Artificial Intelligence and ChatGPT – Academic Integrity [27]
28		Canada	University of Montreal	Montreal Declaration on Responsible AI [28]
29	South America	Canada	McMaster University	Provisional Guidelines on the Use of Generative AI in Teaching and Learning [29]
30		USA	U.S. Department of Education	Artificial Intelligence and the Future of Teaching and Learning [30]
31		Colombia	Universidad del Rosario	Guidelines for the Use of Artificial Intelligence in University Courses [31]
32		Argentina	University of Buenos Aires	Guidelines for the use of ChatGPT and text generative AI in Justice [32]
33		Colombia	Universidad de Los Andes	Guidelines for the use of artificial intelligence in university contexts [33]
34		Colombia	Pontifical Javeriana University	Editorial Policy, Publication Ethics and Malpractice Statement [34]
35		Argentina	Universidad de san andres	Readiness of the judicial sector for artificial intelligence in Latin America [35]

36	Asia	Peru	Government and Digital Transformation Secretariat	National Artificial Intelligence Strategy [36]
37		Chile	Ministry of Science, Technology, Knowledge and Innovation	Guidelines for the use of artificial intelligence tools in the public sector [37]
38		Chile	Pontificia universidad católica de chile	ChatGPT: How to use it in classes? [38]
39		China	Tsinghua University	International AI Cooperation and Governance Forum 2022 [39]
40		Singapore	National University of Singapore	Responsible Use of AI – Guidance from a Singapore Regulatory Perspective [40]
41		Japan	Nagoya University	Regarding the Use of Generative AI [41]
42		Japan	University of Tokyo	Guidelines for Instructors Regarding AI in University Education at Tokyo University of Foreign Studies [42]
43		Hong Kong SAR	University of Hong Kong	Use of Artificial Intelligence Tools in Teaching, Learning and Assessments: A Guide for Students [43]
44		South Korea	Seoul National University (SNU)	Seoul National University AI Policy Initiative [44]
45		Japan	Kyushu University	Note on the Use of Generative AI in Education at Kyushu University – For Teachers [45]
46		Japan	Weseda University	About the Use of Generative Arificial Intelligence (ChatGPT, etc.) [46]
47		South Korea	Ulsan National Institute of Science and Technology	A Guide to the Use of Generative AI [47]
48		Singapore	Sinapore Management University	SUM Framework for the use of Generative AI Tools [48]
49		Singapore	Singapore University of Technology and Design	Artificial Intelligence in Education [49]
50		Singapore	Singapore Institute of Technology	Generative AI at the Singapore Institute of Technology [50]
51		Singapore	Nayang Technological University	NUT Position on the Use of Generative Artificial Intelligence in Research [51]
52		Taiwan	National Tsing Hua University	Guidelines for Collaboration, Co-learning, and Cultivation of Artificial Intelligence Competencies in University Education [52]
53		Taiwan	National Taiwan University	Guidance for Use of Generative AI Tools for Teaching and Learning [53]
54		Hong Kong SAR	The Chinese University of Hong Kong	Use of Artificial Intelligence Tools in Teaching, Learning and Assessments A Guide for Students [54]

55	Europe	Thailand	Chulalongkorn University	Chulalongkorn University Principles and Guidelines for using AI Tools [55]
56		Malaysia	Universiti Malaya	ChatGPT General Usage [56]
57		Malaysia	Universiti Putra Malaysia	Guide for ChatGPT usage in Teaching and Learning [57]
58		China	The Supervision Department of the Ministry of Science and Technology	Guidelines for Responsible Research Conduct (2023) [58]
59		Australia	The Department of Education	The Australian Framework for Generative Artificial Intelligence (AI) in Schools [59]
60		New Zealand	The University of Auckland	Advice for students on using generative artificial intelligence in coursework [60]
61		UK	University of Oxford	Use of generative AI tools to support learning [61]
62		UK	University of Cambridge	Artificial intelligence and teaching, learning, and assessment [62]
63		UK	Imperial College London	Generative AI Guidance [63]
64		UK	London School of Economics and Political Science	School Statement on Generative Artificial Intelligence and Education [64]
65		UK	University College London	Using AI tools in assessment [65]
66		UK	The University of Edinburgh	AI Guidance for Staff and Students [66]
67		Netherlands	Erasmus University Rotterdam	Gnenerative AI in education [67]
68		Belgium	KU Leuven	Responsible use of generative Artificial Intelligence [68]
69		Switzerland	ETH Zurich	AI in education, resources for teaching faculty [69]
70		Netherlands	University of Amsterdam	AI tools and your studies [70]
71		Norway	University of Oslo	Guidelines to use artificial intelligence at UiO [71]
72		Finland	University of Helsinki	Using AI to support learning [72]
73		Italy	University of Padua	Research in Artificial intelligence [73]
74		Sweden	Stockholm University	Guidelines on using AI-powered chatbots in education and research [74]
75		Denmark	Technical University of Denmark	DTU opens up for the use of artificial intelligence in teaching [75]
78		Netherlands	Delft University of Technology	AI chatbots in unsupervised assessment [76]
77		Portugal	Universidade de Lisboa	Artificial Intelligence in education – Técnico presents resolution on the use of tools such as ChatGPT [77]

78		Netherlands	University of Utrecht	Guidelines for the use of generative AI[78]
79		Switzerland	University of Zurich	Guidelines for the Use of AI Tools[79]
80		UK	Russell Group (24 UK research-intensive universities)	Russell Group principles on the use of generative AI tools in education[80]

II. References

- [1] U. of C. Town, “Artificial Intelligence for Teaching & Learning | University of Cape Town.” 2023. [Online]. Available: <https://cilt.uct.ac.za/teaching-resources/artificial-intelligence-teaching-learning>
- [2] C. University, “FCAI Policy and Guidelines for use of Generative AI in Postgraduate Studies and Research,” 2023, [Online]. Available: <https://chat.openai.com/chat>
- [3] U. of Witwatersrand, “ChatGPT & other AI tools for Learning and Teaching.” 2023. [Online]. Available: <https://www.wits.ac.za/media/wits-university/learning-and-teaching/cltd/documents/ChatGPT-for-learning-and-teaching.pdf>
- [4] U. of Pretoria, “Student’s Guide: Leveraging Generative Artificial Intelligence for Learning.” 2023. [Online]. Available: https://www.up.ac.za/media/shared/391/pdfs/up-student-guide_-leveraging-generative-artificial-intelligence-for-learning.zp242396.pdf
- [5] U. of Johannesburg, “UJ PRACTICE NOTE: GENERATIVE ARTIFICIAL INTELLIGENCE IN TEACHING, LEARNING AND RESEARCH.” 2023. [Online]. Available: <https://www.uj.ac.za/wp-content/uploads/2023/08/uj-ai-practice-guide-2023.pdf>
- [6] S. University, “Academic Integrity - Economics - Library guides / Biblioteekgidse at Stellenbosch University.” 2023. [Online]. Available: https://libguides.sun.ac.za/Economics/academic_integrity
- [7] U. of F. State, “Stepping up with ChatGPT - AI-assisted Technology in Education - LibGuides at University of the Free State.” 2023. [Online]. Available: <https://ufs.libguides.com/AI/SteppingupwithChatGPT>
- [8] T. A. U. in Cairo, “AUC’s Statement on the Use of Artificial Intelligence Tools | The American University in Cairo.” 2023. [Online]. Available: <https://www.aucegypt.edu/about/leadership/provost/use-of-artificial-intelligence-tools>
- [9] N.-W. University, “Implications of AI for teaching and learning in higher education & Guidelines for the Utilization of AI in Teaching and Learning at NWU.” 2023. [Online]. Available: <https://services.nwu.ac.za/sites/services.nwu.ac.za/files/files/CoPAI/NWU%20Guidelines%20for%20the%20use%20AI%20for%20Teaching%20and%20Learning.pdf>
- [10] A. O. on R. A. Intelligence, “Generative AI guidelines at South African universities.” 2024. [Online]. Available: <https://www.africanobservatory.ai/social/generative-ai-guidelines-at-south-african-universities>

- [11] H. University, “Guidelines for Using ChatGPT and other Generative AI tools at Harvard | Office of the Provost.” 2023. [Online]. Available: <https://provost.harvard.edu/guidelines-using-chatgpt-and-other-generative-ai-tools-harvard>
- [12] S. University, “Generative AI Policy Guidance | Office of Community Standards.” 2023. [Online]. Available: <https://communitystandards.stanford.edu/generative-ai-policy-guidance>
- [13] M. I. T, “Getting Started with AI-Enhanced Teaching: A Practical Guide for Instructors - MIT Sloan Teaching & Learning Technologies.” 2023. [Online]. Available: <https://mitsloanedtech.mit.edu/ai/teach/getting-started/>
- [14] P. University, “Generative AI | McGraw Center for Teaching and Learning.” 2023. [Online]. Available: <https://mcgraw.princeton.edu/generative-ai>
- [15] U. of Chicago, “Guidance for Syllabus Statements on the Use of AI Tools.” 2023. [Online]. Available: https://teaching.uchicago.edu/sites/default/files/2023-09/CCTL_AI%20Syllabus%20Statements.pdf
- [16] C. University, “Considerations for AI Tools in the Classroom.” 2023. [Online]. Available: <https://ctl.columbia.edu/resources-and-technology/resources/ai-tools/>
- [17] C. I. of Technology, “Generative AI - Information Management Systems and Services.” 2023. [Online]. Available: <https://www.imss.caltech.edu/services/ai>
- [18] B. University of California, “Appropriate Use of ChatGPT and Similar AI Tools | Office of Ethics.” 2023. [Online]. Available: <https://ethics.berkeley.edu/privacy/appropriate-use-chatgpt-and-similar-ai-tools>
- [19] Y. University, “Guidelines for the Use of Generative AI Tools | Office of the Provost.” 2023. [Online]. Available: <https://provost.yale.edu/news/guidelines-use-generative-ai-tools>
- [20] U. of Pennsylvania, “Statement on Guidance for the University of Pennsylvania Community on Use of Generative Artificial Intelligence | UPenn ISC.” 2023. [Online]. Available: <https://www.isc.upenn.edu/security/AI-guidance>
- [21] L. A. University of California, “Guidance for the Use of Generative AI – UCLA Center for the Advancement of Teaching.” 2023. [Online]. Available: https://teaching.ucla.edu/resources/ai_guidance/
- [22] C. University, “Artificial Intelligence (AI) | IT@Cornell.” 2023. [Online]. Available: <https://it.cornell.edu/ai>
- [23] U. of Toronto, “Generative Artificial Intelligence in the Classroom: FAQ’s – Office of the Vice-Provost, Innovations in Undergraduate Education.” 2024. [Online]. Available: <https://www.viceprovostundergrad.utoronto.ca/16072-2/teaching-initiatives/generative-artificial-intelligence/>
- [24] U. of B. Columbia, “UBC Guidance - Generative AI.” 2023. [Online]. Available: <https://genai.ubc.ca/guidance/>
- [25] M. University, “Principles on Generative AI in Teaching and Learning at McGill.” 2023. [Online]. Available:

https://www.mcgill.ca/provost/files/provost/principles_on_generative_ai_in_teaching_and_learning_at_mcgill.pdf

- [26] U. of Alberta, “AI-Squared - Artificial Intelligence and Academic Integrity | Centre for Teaching and Learning.” 2023. [Online]. Available: <https://www.ualberta.ca/centre-for-teaching-and-learning/teaching-toolkit/teaching-in-the-context-of-ai/artificial-intelligence-academic-integrity.html>
- [27] U. of Waterloo, “Artificial intelligence and ChatGPT | Academic Integrity.” 2023. [Online]. Available: <https://uwaterloo.ca/academic-integrity/artificial-intelligence-and-chatgpt>
- [28] U. of Montreal, “About - Déclaration de Montréal IA responsable.” 2017. [Online]. Available: <https://montrealdeclaration-responsibleai.com/about/>
- [29] M. University, “Provisional Guidelines on the Use of Generative AI in Teaching and Learning.” 2023. [Online]. Available: <https://provost.mcmaster.ca/office-of-the-provost-2/generative-artificial-intelligence/task-force-on-generative-ai-in-teaching-and-learning/provisional-guidelines-on-the-use-of-generative-ai-in-teaching-and-learning/>
- [30] U. S. D. of Education, “Artificial Intelligence and the Future of Teaching and Learning.” 2023. [Online]. Available: <https://www2.ed.gov/documents/ai-report/ai-report.pdf>
- [31] U. del Rosario, “Guidelines for the Use of Artificial Intelligence in University Courses.” 2023. [Online]. Available: <https://forogpp.files.wordpress.com/2023/02/guidelines-for-the-use-of-artificial-intelligence-in-university-courses-v4.3.pdf>
- [32] U. of B. Aires, “Guidelines for the use of ChatGPT and text generative AI in Justice.” 2023. [Online]. Available: <https://ialab.com.ar/wp-content/uploads/2024/02/Guia-uso-IAG-.pdf>
- [33] U. de L. Andes, “Guidelines for the use of artificial intelligence in university contexts.” 2023. [Online]. Available: <https://juangutierrezco.files.wordpress.com/2023/08/guidelines-for-the-use-of-artificial-intelligence-in-university-contexts-v5.0.pdf>
- [34] P. J. University, “Editorial Pontificia Universidad Javeriana. Editorial Policy, Publication Ethics and Malpractice Statement.” 2023. [Online]. Available: <https://revistas.javeriana.edu.co/files-articulos/CRC-EPUJ/manuales/ETHICS/>
- [35] U. de san andres, “COMPILED: READINESS OF THE JUDICIAL SECTOR FOR ARTIFICIAL INTELLIGENCE IN LATIN AMERICA.” 2023. [Online]. Available: <https://cetys.lat/wp-content/uploads/2021/09/compilado-eng.pdf>
- [36] G. and D. T. Secretariat, “NATIONAL ARTIFICIAL INTELLIGENCE STRATEGY.” 2021. [Online]. Available: <https://cdn.www.gob.pe/uploads/document/file/1909267/National%20Artificial%20Intelligence%20Strategy%20-%20Peru.pdf>
- [37] T. Ministry of Science, “Guidelines for the use of artificial intelligence tools in the public sector.” 2023. [Online]. Available: https://minciencia.gob.cl/uploads/filer_public/ae/9a/ae9a7ce7-807b-4781-9ac3-9b253bfbe735/of_n711_2023_dis_lin_ia_minciencia.pdf
- [38] pontificia universidad católica de chile, “ChatGPT: How to use it in classes?” 2023. [Online]. Available: <https://www.uc.cl/noticias/chatgpt-como-usarlo-en-clases/>

- [39] T. University, “About the Forum-INSTITUTE FOR AI INTERNATIONAL GOVERNANCE TSINGHUA UNIVERSITY.” 2022. [Online]. Available: https://aiig.tsinghua.edu.cn/en/International_Forum/2022/About_the_Forum.htm
- [40] N. U. of Singapore, “B_Oct_Responsible_Use_of_AI - Centre for Technology, Robotics, Artificial Intelligence and the Law.” 2023. [Online]. Available: <https://law.nus.edu.sg/trail/responsible-use-of-ai/>
- [41] N. University, “Nagoya University Academics / Campus Life.” 2023. [Online]. Available: <https://en.nagoya-u.ac.jp/academics/ai/index.html>
- [42] U. of Tokyo, “Guidelines for Instructors Regarding AI in University Education at Tokyo University of Foreign Studies.” 2023. [Online]. Available: https://www.tufs.ac.jp/documents/education/guideline/ai_guideline_en.pdf
- [43] U. of H. Kong, “Use of Artificial Intelligence Tools in Teaching, Learning and Assessments A Guide for Students.” 2023. [Online]. Available: https://www.aqs.cuhk.edu.hk/documents/A-guide-for-students_use-of-AI-tools.pdf
- [44] S. N. U. (SNU), “Seoul National University AI Policy Initiative.” 2020. [Online]. Available: https://sapi.co.kr/wp-content/uploads/2021/04/SAPI-%ED%99%9C%EB%8F%99%EB%B3%B4%EA%B3%A0%EC%84%9C_ENG.pdf
- [45] K. University, “Note on the Use of Generative AI in Education at Kyushu University – For Teachers -.” 2023. [Online]. Available: <https://www.kyushu-u.ac.jp/en/notices/view/461/>
- [46] W. University, “About the Use of Generative Artificial Intelligence (ChatGPT, etc.) – Waseda University.” 2023. [Online]. Available: <https://www.waseda.jp/top/en/news/77786>
- [47] U. N. I. of S. and Technology, “A Guide to the Use of Generative AI.” 2023. [Online]. Available: <https://heyzine.com/flip-book/3a9d4cb37e.html#page/1>
- [48] SMU, “SMU Framework for The Use of Generative AI Tools,” <https://cte.smu.edu.sg/resources/smu-framework-generative-ai>.
- [49] S. U. of T. and Design, “Singapore University of Technology and Design (SUTD): campusX.” 2023. [Online]. Available: <https://www.sutd.edu.sg/campusX/AI-in-Education>
- [50] S. I. of Technology, “Generative AI at the Singapore Institute of Technology (SIT) – Teaching Connections.” 2023. [Online]. Available: <https://blog.nus.edu.sg/teachingconnections/2024/01/24/generative-ai-at-the-singapore-institute-of-technology-sit/>
- [51] N. T. University, “NTU Position on the Use of Generative Artificial Intelligence (GAI) in Research | Research | NTU Singapore.” 2023. [Online]. Available: <https://www.ntu.edu.sg/research/resources/use-of-gai-in-research>
- [52] N. T. H. University, “Guidelines for Collaboration, Co-learning, and Cultivation of Artificial Intelligence Competencies in University Education.” 2023. [Online]. Available: <http://ctld.nthu.edu.tw/var/file/217/1217/img/555612445.pdf>
- [53] National Taiwan University, “Guidance for Use of Generative AI Tools for Teaching and Learning,” <https://www.dlc.ntu.edu.tw/en/ai-tools-en/>.

- [54] T. C. U. of H. Kong, “Use of Artificial Intelligence Tools in Teaching, Learning and Assessments A Guide for Students.” 2023. [Online]. Available: https://www.aqs.cuhk.edu.hk/documents/A-guide-for-students_use-of-AI-tools.pdf
- [55] C. University, “Chulalongkorn University Principles and Guidelines for using AI Tools – Chulalongkorn University.” 2023. [Online]. Available: <https://www.chula.ac.th/en/news/125190/>
- [56] U. Malaya, “CHATGPT GENERAL USAGE.” 2023. [Online]. Available: https://asdpd.um.edu.my/doc/E-Learning/Newsletter_Series%202_2023_ChatGPT_Usage.pdf
- [57] U. P. Malaysia, “GUIDE FOR CHATGPT USAGE IN TEACHING AND LEARNING.” 2023. [Online]. Available: https://cade.upm.edu.my/upload/dokumen/20230202105701Guide_for_ChatGPT_Usage_in_Teaching_and_Learning.pdf
- [58] T. S. D. of the M. of S. and Technology, “Guidelines for Responsible Research Conduct.” 2023. [Online]. Available: <https://www.most.gov.cn/kjbgz/202312/W020231221582942330036.pdf>
- [59] T. D. of Education, “The Australian Framework for Generative Artificial Intelligence (AI) in Schools - Department of Education, Australian Government.” 2023. [Online]. Available: <https://www.education.gov.au/schooling/announcements/australian-framework-generative-artificial-intelligence-ai-schools>
- [60] T. U. of Auckland, “Advice for students on using Generative Artificial Intelligence in coursework - The University of Auckland.” 2023. [Online]. Available: <https://www.auckland.ac.nz/en/students/forms-policies-and-guidelines/student-policies-and-guidelines/academic-integrity-copyright/advice-for-student-on-using-generative-ai.html>
- [61] U. of Oxford, “Use of generative AI tools to support learning | University of Oxford.” 2023. [Online]. Available: <https://www.ox.ac.uk/students/academic/guidance/skills/ai-study>
- [62] U. of Cambridge, “Artificial intelligence and teaching, learning and assessment.” 2023. [Online]. Available: <https://www.cambridgeinternational.org/support-and-training-for-schools/artificial-intelligence/>
- [63] I. C. London, “Generative AI guidance | Administration and support services | Imperial College London.” 2023. [Online]. Available: <https://www.imperial.ac.uk/admin-services/library/learning-support/generative-ai-guidance/>
- [64] L. S. of E. and P. Science, “School statement on Generative Artificial Intelligence and education.” 2023. [Online]. Available: <https://info.lse.ac.uk/staff/divisions/Eden-Centre/Assets-EC/Documents/AI-web-expansion-Sept-23/Updated-School-statement-on-Artificial-Intelligence-Nov-10-2023-FINAL.pdf>
- [65] U. C. London, “Using AI tools in assessment | Teaching & Learning - UCL – University College London.” 2023. [Online]. Available: <https://www.ucl.ac.uk/teaching-learning/generative-ai-hub/using-ai-tools-assessment>
- [66] T. U. of Edinburgh, “Generative Artificial Intelligence guidance for students | The University of Edinburgh.” 2023. [Online]. Available: <https://www.ed.ac.uk/bayes/ai-guidance-for-staff-and-students/ai-guidance-for-students>

- [67] E. U. Rotterdam, “AI tools in education | Erasmus University Rotterdam.” 2023. [Online]. Available: <https://www.eur.nl/en/about-eur/vision/community-learning-and-innovation/ai-tools-education>
- [68] K. U. Leuven, “Responsible use of generative Artificial Intelligence.” 2023. [Online]. Available: <https://www.kuleuven.be/english/genai>
- [69] E. T. H. Zurich, “AI in education | ETH Zurich.” 2024. [Online]. Available: <https://ethz.ch/en/the-eth-zurich/education/ai-in-education.html>
- [70] U. of Amsterdam, “AI tools and your studies - student.uva.nl.” 2024. [Online]. Available: <https://student.uva.nl/en/topics/ai-tools-and-your-studies>
- [71] U. of Oslo, “Artificial intelligence (AI) at UiO - University of Oslo.” 2023. [Online]. Available: <https://www.uio.no/english/services/ai/>
- [72] U. of Helsinki, “Using AI to support learning | Instructions | Studies Service | University of Helsinki.” 2023. [Online]. Available: <https://studies.helsinki.fi/instructions/article/using-ai-support-learning>
- [73] U. of Padua, “Artificial Intelligence | DEI.” 2023. [Online]. Available: <https://www.dei.unipd.it/en/computer-engineering/artificial-intelligence>
- [74] S. University, “Guidelines on using AI-powered chatbots in education and research - Staff.” 2023. [Online]. Available: <https://www.su.se/staff/services/teaching/guidelines-on-using-ai-powered-chatbots-in-education-and-research-1.649009>
- [75] T. U. of Denmark, “DTU opens up for the use of artificial intelligence in teaching.” 2024. [Online]. Available: <https://www.dtu.dk/english/newsarchive/2024/01/dtu-opens-up-for-the-use-of-artificial-intelligence-in-teaching>
- [76] UT Delft, “AI chatbots in unsupervised assessment ,” <https://www.tudelft.nl/en/teaching-support/didactics/assess/guidelines/ai-chatbots-in-projects-and-assignments>, 2023.
- [77] U. de Lisboa, “Artificial Intelligence in education – Técnico presents resolution on the use of tools such as ChatGPT – Técnico Lisboa.” 2024. [Online]. Available: <https://tecnico.ulisboa.pt/en/news/campus-community/artificial-intelligence-in-education-tecnico-presents-resolution-on-the-use-of-tools-such-as-chatgpt/>
- [78] U. of Utrecht, “Guidelines for the use of generative AI - Corporate Identity - Utrecht University.” 2023. [Online]. Available: <https://www.uu.nl/en/organisation/corporate-identity/guidelines-for-the-use-of-generative-ai>
- [79] university of Zurich, “Guidelines for the Use of AI Tools,” <file:///Users/salehafroogh/Downloads/Guidelines%20on%20ChatGPT%20and%20AI-4.pdf>, 2023.
- [80] R. G. (24 U. K. research-intensive universities), “Russell Group principles on the use of generative AI tools in education.” 2023. [Online]. Available: https://russellgroup.ac.uk/media/6137/rg_ai_principles-final.pdf

III. Texts of 80 Guidelines with a Detailed Table of Contents

Table of Contents

1. Artificial Intelligence for Teaching & Learning	30
1.1. Staff Guide: Assessment and academic integrity in the age of AI	30
Communicating with your students	31
Academic integrity strategies	33
How students use generative AI tools for support.....	33
AI detection tools.....	34
Risks and concerns of using AI.....	35
Designing assessments for higher order thinking	36
Using prompts for preparation	38
1.2. Staff Guide: Teaching and learning with AI tools	40
Introduction	40
Setting expectations	40
Ethics of Using ChatGPT (or similar tools) in education.....	41
Important literacies associated with AI.....	42
AI Literacy	43
Critical literacy	44
Information literacy.....	45
AI and academic integrity.....	45
Roles and uses to explore for AI tools in teaching and learning	46
What AI tools are there for education?	48
1.3. Student Guide: Using ChatGPT and other AI tools in education.....	48
What is AI?	49
How does ChatGPT generate its outputs?	50
Risks and Ethics of using ChatGPT (or similar tools) in education	50
How ChatGPT can help you in your learning processes?	51
AI Tools and Academic Integrity	52
The UCT Code of Conduct	54
1.4. Student Guide: Developing effective prompts for generative AI tools	54
Understanding how generative AI works	55
Prompts 'cheat sheet'	55
Useful prompt patterns strategies.....	56
Patterns based on inputs and outputs	57
Patterns based on refining and interrogating the questions	59
Patterns which use a standard template	61
Take away points	62

1.5. Student Guide: Ethical use of generative AI for research purposes	84
General use.....	84
General characteristics.....	85
Research integrity implications	87
Research process stages with generative AI examples	88
Stage of research: Conceptualisation and research idea generation	89
Stage of research: Literature review	91
Stage of research: Data collection/transcription	94
Stage of research: Data analysis	94
Stage of research: Write up.....	96
Stage of research: Dissemination.....	97
Bespoke research tools.....	98
Conclusion	99
2. Policy and Guidelines for use of Generative AI.....	100
Introduction.....	100
Policy Development Approach	101
Key Guidelines for Using Generative AI in Teaching and Research.....	101
Scope of the Guidelines.....	101
Coursework	102
Research Proposals, Thesis, and Dissertations	102
Referencing Guidelines	104
3. ChatGPT & other AI tools for Learning and Teaching.....	104
Introduction to ChatGPT	104
What is ChatGPT?.....	105
Relevant resources.....	105
How ChatGPT works	105
Opportunities for learning & teaching offered by ChatGPT	107
Ethics and morality of using ChatGPT for Learning and Teaching.....	109
How ChatGPT can be used in the classroom	110
Lesson Plans for Using ChatGPT in Higher Education	114
Assessing student learning	115
Misuse of ChatGPT by students / Academic Integrity and ChatGPT	117
Detection of misuse of ChatGPT by students	119
Policy and guidelines	120
Annexure: AI Tools for Learning and Teaching.....	121
Tutorials, Courses and Explanatory videos	126
Prompt Development/Engineering	127
Research Tools.....	127

4. Leveraging Generative Artificial Intelligence for Teaching and Learning Enhancement.....	129
General Principles and Guidelines.....	130
Academic Integrity	132
Plagiarism.....	132
Strategies for Optimising AI Prompts and Achieving Desired Results	132
Use Generative AI to Enhanced Learning Task Examples of prompts (input).....	133
Boosting Student Productivity with Generative AI	134
Providing Personalized 24/7 Tutor Support via Generative AI.....	134
Providing Support	135
Summary.....	136
5. UJ Practice Notes: Generative Artificial Intelligence.....	136
BENEFITS OF GENERATIVE AI TOOLS FOR LEARNING, TEACHING AND RESEARCH	137
TEACHING, LEARNING AND RESEARCH INTEGRITY.....	141
ASSESSMENT INTEGRITY	143
ACADEMIC INTEGRITY	144
ETHICAL CONSIDERATIONS IN THE USE OF GENERATIVE AI	145
CONCLUSION	146
RECOMMENDATIONS.....	146
6. Stellenbosch University Academic Integrity: Responsible Use of AI tools	147
Description.....	147
Are students allowed to use ChatGPT and other AI tools?	148
Responsible use of AI tools	148
Is it safe to use ChatGPT for your assignment/essay?	149
Declaration example	149
7. Stepping up with ChatGPT - AI-assisted Technology in Education	150
How and why is ChatGPT relevant for academics?	150
Getting started with ChatGPT.....	151
Applications of ChatGPT for academics.....	151
Tips for using ChatGPT effectively	152
Plagiarism.....	155
How to spot ChatGPT-generated text.....	155

8. AUC's Statement on the Use of Artificial Intelligence Tool	155
9. Guidelines for the Utilization of AI in Teaching and Learning at NWU.....	156
Continuity and Change.....	157
Higher Education	158
Selected affordances linked to the responsible use of AI	158
Selected perceived risks associated with AI	161
An emerging NWU perspective on AI and the Curriculum	162
Academic Integrity and accountability for the uses of AI.....	163
NWU Guidelines as regards the place of AI in the Curriculum and NWU Values	163
10. Generative AI guidelines at South African universities	165
11. Guidelines for Using ChatGPT and other Generative AI tools at Harvard.....	169
12. Generative AI Policy Guidance	170
Honor Code Implications of Generative AI Tools.....	170
13. Getting Started with AI-Enhanced Teaching.....	171
The Basics.....	172
Generative AI Tools.....	173
Ethical Considerations.....	174
Data Privacy.....	174
Falsehoods and Bias	175
AI-Powered Teaching Strategies.....	175
1. Use AI to Generate Concrete Examples	176
2. Use AI to Create Practice Quizzes.....	177
3. Assign Students to Generate Visual Summaries	178
4. Ask Students to Teach the AI.....	179
Get Support	180
Conclusion	180
14. Generative AI Guidance.....	180
Syllabus Language	181
Ethical and Other Risks	181
Equity and Access	181
Student data and privacy	182
Inaccuracies and fabrication	182
Cognitive Offloading	182
Bias and stereotypes	183
Labor concerns with how AI tools are trained	183
Environmental Impact.....	183
Implications for Assignments	183
Assigning Generative AI.....	185

Detection Software and Academic Integrity	186
15. Guidance for Syllabus Statements on the Use of AI Tools.....	188
Considerations and Examples of Syllabus Statements on the Use of AI Tools.....	188
16. Considerations for AI Tools in the Classroom	191
Considerations for AI Tools in the Classroom.....	191
For instructors: share your thoughts with us!	191
What are AI Tools?	191
ChatGPT	192
ChatGPT in the Classroom	193
Develop Course Policies that Include Digital Transparency	194
Scaffold Activities and Assignments	194
Design Authentic Assessments for Learning	195
Incorporate AI Tools into Assignment Design	195
Conclusion	196
The CTL is here to help!.....	196
Columbia CTL Related Resources	197
17. Guidance on the Use of Generative AI and Large Language Model Tools	197
Guidance on the Use of Generative AI and Large Language Model Tools	197
18. Appropriate use of ChatGPT and Similar AI Tools.....	199
Allowable Use:.....	199
Prohibited Use:	200
Additional Guidance:.....	200
Rationale for the Above Guidance:.....	200
Guidance on Appropriate Use	201
References	201
Additional Readings	202
19. Guidelines for the Use of Generative AI Tools	202
20. Statement on Guidance for the University of Pennsylvania Community.....	204
Scope.....	204
General Guidance for Penn Community (Educators, Staff, Researchers, and Students)	205
21. Guidance for the use of generative AI.....	209
About this guide	209

What is ChatGPT?	209
Are students using ChatGPT?	210
Explore how ChatGPT works	210
Try out ChatGPT and Reflect: What are the opportunities?	210
Recommendation.....	210
Ideas for Updating Your Course Activities with AI in Mind	212
Communicate: Talk With Your Students About AI	213
Examples and Resources	217
22. Cornell Guidelines for artificial intelligence.....	218
Guidelines.....	218
Accountability.....	218
Confidentiality and Privacy.....	219
Use for Education and Pedagogy.....	219
Use for Research.....	219
Use for Administration and Other Purposes	219
Cornell AI Services and Pilots	219
Recommended Resources	220
University Task Force Reports.....	220
AI Questions, Concerns, Needs, or Ideas?.....	220
Comments?.....	220
23. Generative Artificial Intelligence in the classroom.....	220
Sample Syllabus Statements.....	221
SGS Guidelines	221
Copyright Considerations.....	221
Frequently Asked Questions	222
24. Generative AI – Academic Integrity at UBC.....	225
Thinking about ChatGPT?	225
What is ChatGPT and what is generative artificial intelligence?	225
Can instructors test AI tools to see what they can do, and what they cannot?	225
Is the use of AI tools considered to be academic misconduct at UBC?.....	226
How can instructors use AI tools in their classroom?.....	227
Can students use AI tools to complete assignments?	227
Can or should instructors use AI detectors to detect the use of AI tools on assignments or assessments?	228
Can I use Turnitin’s new AI-detection feature in my courses?	228
What should I do if I suspect the unauthorized use of AI tools?	229
What are the recommendations for citing content developed by Generative AI if its use is permitted?.....	229
Is there funding available for teaching and learning projects on generative AI?	230
25. Principles on Generative AI in Teaching and Learning at McGill.....	230

26. AI-Squared – Artificial Intelligence and Academic Integrity	230
Dialogue with students.....	234
27. Artificial Intelligence and ChatGPT – Academic Integrity	235
INSTRUCTOR RESOURCES FOR ARTIFICIAL INTELLIGENCE & CHATGPT TOOLS.....	236
TIPS FOR STUDENTS.....	236
28. Montreal Declaration on Responsible AI.....	238
Our process for responsible artificial intelligence	238
29. Provisional Guidelines on the Use of Generative AI in Teaching and Learning	239
Provisional Principles.....	239
Provisional Guidelines on the Use of Generative AI in Teaching and Learning.....	240
General Guidelines	240
Instructor or TA Use of Generative AI	241
Privacy, Security and Selection of Tools.....	242
Student Assessments	243
Sample Syllabus Statements.....	243
Use Prohibited.....	244
Some Use Permitted	244
30. Artificial Intelligence and the Future of Teaching and Learning	245
Introduction.....	245
Rising Interest in AI in Education	246
Three Reasons to Address AI in Education Now	247
Toward Policies for AI in Education	248
Building Ethical, Equitable Policies Together	249
Guiding Questions.....	250
Foundation 1: Center people (parents, educators, and students).....	250
Foundation 2: Advance Equity	251
Foundation 3: Ensure Safety, Ethics, and Effectiveness.....	252
Foundation 4: Promote Transparency	253
Overview of Document.....	253
What is AI?	254
Perspective: Human-Like Reasoning	254
Perspective: An Algorithm that Pursues a Goal	255
Perspective: Intelligence Augmentation.....	256
Definition of “Model”	257
Insight: AI Systems Enable New Forms of Interaction	258
Key Recommendation: Human in the Loop AI.....	258
Learning	259
Insight: AI Enables Adaptivity in Learning.....	260
Intelligent Tutoring Systems: An Example of AI Models	261
Important Directions for Expanding AI-Based Adaptivity.....	262
A Duality: Learning With and About AI.....	264

A Challenge: Systems Thinking About AI in Education	264
Open Questions About AI for Learning	265
Key Recommendation: Seek AI Models Aligned to a Vision for Learning	266
Teaching.....	266
Always Center Educators in Instructional Loops.....	267
Insight: Using AI to Improve Teaching Jobs.....	267
Preparing and Supporting Teachers in Planning and Reflecting.....	269
Designing, Selecting, and Evaluating AI Tools	270
Challenge: Balancing Human and Computer Decision-Making	270
Challenge: Making Teaching Jobs Easier While Avoiding Surveillance	271
Challenge: Responding to Students' Strengths While Protecting Their Privacy	272
Questions Worth Asking About AI for Teaching.....	273
Key Recommendation: Inspectable, Explainable, Overridable AI	274
Formative Assessment.....	275
Building on Best Practices.....	275
Implications for Teaching and Learning.....	276
Insight: AI Can Enhance Feedback Loops.....	277
An Example: Automated Essay Scoring.....	278
Key Opportunities for AI in Formative Assessment.....	279
Key Recommendation: Harness Assessment Expertise to Reduce Bias.....	280
Related Questions.....	280
Research and Development.....	281
Insight: Research Can Strengthen the Role of Context in AI	281
Attention to the Long Tail of Learner Variability	283
Partnership in Design-Based Research.....	284
Re-thinking Teacher Professional Development.....	285
Connecting with Public Policy	286
Key Recommendation: Focus R&D on Addressing Context	287
Ongoing Questions for Researchers.....	287
Desired National R&D Objectives.....	288
Recommendations.....	288
Insight: Aligning AI to Policy Objectives.....	289
Calling Education Leaders to Action.....	290
Recommendation #1: Emphasize Humans in the Loop	290
Recommendation #2: Align AI Models to a Shared Vision for Education	291
Recommendation #3: Design Using Modern Learning Principles	293
Recommendation #4: Prioritize Strengthening Trust	294
Recommendation #5: Inform and Involve Educators	295
Recommendation #6: Focus R&D on Addressing Context and Enhancing Trust and Safety	296
Recommendation #7: Develop Education-Specific Guidelines and Guardrails.....	297
Next Steps.....	298
31. Guidelines for the Use of Artificial Intelligence in University Courses.....	299
1. Justification:	299
2. Objectives:	300
3. Rules for the use of AI in and out of the classroom:	300
4. Why is an informed, transparent, ethical, and responsible use of AI necessary?	301

5. Open-Guidelines:	303
32. Guidelines for the use of ChatGPT and text generative AI in Justice.....	303
I. Executive Summary	303
II. Goal	305
III. Conceptual framework	305
IV. Possible uses of text generative AI in Justice	307
V. Relevant considerations regarding the potential uses of text generative AI	308
VI. Specific guidelines and recommendations for the use of AI in Justice	309
VII. Team:.....	319
VIII. Do you want to be part of the Guide?	319
33. Guidelines for the use of artificial intelligence in university contexts.....	320
Introduction.....	320
Objectives of these guidelines	321
Rules for the use of AI in and out of the classroom.....	322
Why the informed, transparent, ethical, and responsible use of AI in academia is necessary?	324
Other resources for teachers	326
Guidelines open to change	327
Changes introduced in this version of the guideline	327
Use of these guidelines in your classes.....	328
34. Editorial Policy, Publication Ethics and Malpractice Statement	328
Editorial Policy	328
Copyright and Licensing.....	330
Privacy Statement.....	330
No Author Fees.....	331
Publication Ethics Statement	332
COPE adherence Statement.....	333
Double-Blind Peer-Review Process	342
Open Access and Self Archiving Policy	343
Advertising and Marketing	343
35. Readiness of the judicial sector for artificial intelligence in Latin America.....	344
Justice and Artificial Intelligence:	344

History of AI updating, digitalization and inclusion, and links with the judicial sector	346
An exploratory analytical framework designed for the Latin American reality	350
Mapping of similarities and differences between national cases	361
36. National Artificial Intelligence Strategy	364
Artificial Intelligence.....	364
Definition for the Strategy.....	364
Machine Learning (Automatic or machine learning).....	365
External Context.....	365
External Analysis: Use Cases of Local Government.....	366
External Context: Use Cases in Justice:	366
External Context: Use Case in Citizen Security	367
External Context: Use Case in Education.....	367
External Analysis: Use Cases in Agriculture	367
Use Cases in Fisheries and Aquaculture.....	368
Use Cases in Forest Protection:	368
Use Cases in Telecommunications.....	368
Use Cases for the Mining Industry:.....	368
Use Cases in Energy:	369
Ethics in AI:	369
Internal Analysis:.....	369
Main organizations using AI in Peru:	370
Labor market situation in Peru Photo:.....	370
State of the local academy in terms of training of professionals with AI knowledge:	370
Results of the Peruvian Public Sector AI Status Survey (January - 2021) Question:.....	371
Applications with AI Components Developed During the COVID Pandemic in Peru:	371
Purpose of the National Strategy for Artificial Intelligence:	373
Strategic objectives:.....	374
E1 - Training and Talent Attraction.....	374
E2 - Economic model	375
E3 - Technological infrastructure.....	377
E4 - Data	378
E5 - Ethics.....	379

NATIONAL CENTER FOR INNOVATION AND ARTIFICIAL INTELLIGENCE.....	380
NATIONAL HIGH PERFORMANCE COMPUTER CENTER	381
<i>37. Guidelines for the use of artificial intelligence tools in the public sector</i>	<i>381</i>
BACKGROUND	381
RECOMMENDED GUIDELINES	382
<i>38. ChatGPT: How to use it in classes?.....</i>	<i>385</i>
A threat?	386
How to use ChatGPT in classes?.....	387
What does the UC propose?	388
What is the future of ChatGPT?.....	389
<i>39. International AI Cooperation and Governance Forum 2022</i>	<i>389</i>
Background and Objective	391
<i>40. Responsible Use of AI – Guidance from a Singapore Regulatory Perspective.....</i>	<i>393</i>
I. Introduction	393
II. AI Verify.....	394
III. Advisory Guidelines	395
IV. Responsible use of AI in the financial sector	395
V. Concluding Remarks.....	397
<i>41. Regarding the Use of Generative AI</i>	<i>397</i>
Regarding the Use of Generative AI	397
<i>42. Guidelines for Instructors Regarding AI in University Education.....</i>	<i>398</i>
How AI Generates Text AI text generators work based on large language models (LLM)	398
Limitations of AI Text Generation	399
Guidelines for the use of AI in university education	399
Resources for Understanding and Using Article Generation AI.....	400
<i>43. Use of Artificial Intelligence Tools in Teaching, Learning and Assessments</i>	<i>401</i>
<i>44. Seoul National University AI Policy Initiative</i>	<i>405</i>
<i>45. Note on the Use of Generative AI in Education at Kyushu University.....</i>	<i>410</i>
<i>46. About the Use of Generative Arificial Intelligence (ChatGPT, etc.).....</i>	<i>413</i>
About the Use of Generative Artificial Intelligence (ChatGPT, etc.)	413
Intellectual resilience and flexible sensitivity.....	413
Fundamental Attitudes towards Generative AI	414
Characteristics of Generative AI.....	414

Usage and Limitations of Generative AI.....	416
Future Action	417
47. A Guide to the Use of Generative AI	417
Preface	418
What is Generative AI?.....	418
Guidelines on Using Generative AI For Faculty	421
Guidelines on Using Generative AI For Students	422
Guidelines on Using Generative AI For Researchers.....	423
Tips for Using Generative AI.....	424
48. SUM Framework for the use of Generative AI Tools.....	431
49. Artificial Intelligence in Education	436
Concluding Statement.....	437
50. Generative AI at the Singapore Institute of Technology.....	438
A Collaborative Approach to Professional Development	438
Guarding the Validity of Assessments.....	439
Concluding Remarks, as Generated by ChatGPT	440
51. NUT Position on the Use of Generative Artificial Intelligence in Research.....	440
NTU Position on the Use of Generative Artificial Intelligence in Research.....	440
52. Guidelines for Collaboration, Co-learning, and Cultivation of AI	443
Students and educators should thoroughly understand the following challenges when using AI:.....	444
Integrating AI in Teaching and Learning:.....	444
Cultivating AI Competence	445
53. Guidance for Use of Generative AI Tools for Teaching and Learning.....	447
Guidance for Instructors	448
Guidance for Students	450
54. Use of Artificial Intelligence Tools.....	452
55. Chulalongkorn University Principles and Guidelines for using AI Tools.....	456
Principles.....	456
Guidelines.....	457
56. ChatGPT General Usage	457
57. Guide for ChatGPT usage in Teaching and Learning.....	458
58. Guidelines for Responsible Research Conduct (2023).....	459

1. Research Topic Selection and Implementation.....	460
(1) Scientific researchers.....	460
(2) Scientific research units.....	461
2. Data management.....	462
(1) Scientific researchers.....	462
(2) Scientific research units.....	467
(3) Academic publishing units.....	467
3. Reviewer.....	469
(1) Scientific researchers.....	471
(2) Scientific research units.....	472
4 . Academic exchanges and cooperation	473
(1) Scientific researchers.....	474
(2) Scientific research units.....	475
(3) Scientific research units.....	476
5. Supervision and management.....	477
(1) Scientific research units.....	477
(2) Scientific research funding agencies.....	477
(3) Scientific researchers.....	478
59. <i>The Australian Framework for Generative Artificial Intelligence (AI) in Schools</i> .	479
Teaching and Learning Generative	481
Human and Social Wellbeing Generative	481
Transparency School communities understand how generative.....	482
Fairness Generative	482
Accountability Generative.....	482
Privacy, Security and Safety Students and others using generative.....	482
60. <i>Advice for students on using generative artificial intelligence in coursework.....</i>	483
What is Generative Artificial Intelligence or Gen-AI?	483
How and when you can use Gen-AI?	483
Risks to using Gen-AI.....	485
How Gen-AI may support your studies	486
Misusing Gen-AI.....	487
Microsoft Copilot.....	487
Additional information.....	488
If you're not sure about using Gen-AI	488
61. <i>Use of generative AI tools to support learning.....</i>	488
Overview	488
Ethical use of generative AI tools.....	489

Five things to think about when using generative AI tools	489
Six tips to keep in mind when using generative AI tools	490
Five ideas for academic reading.....	490
Suggestions of prompts to try:.....	492
Five ideas for academic writing and presentation skills	492
Suggestions of prompts to try:.....	493
Five ideas for supporting your learning	494
Suggestions of prompts to try:.....	495
Five tips for selecting the right generative AI tools for the task	496
Further resources.....	497
62. Artificial intelligence and teaching, learning, and assessment.....	498
Artificial intelligence and teaching, learning and assessment.....	498
63. Generative AI Guidance.....	499
What are Generative AI tools?	499
Acknowledging and referencing and generative AI tools.....	500
Considerations for use of generative AI tools	501
64. School Statement on Generative Artificial Intelligence and Education.....	503
65. Engaging with AI in your education and assessment.....	504
Using AI tools in assessment	504
Introduction.....	504
Using AI to support learning	504
Keeping up with developments	505
Guidance	505
The categories of assessment.....	505
1. AI tools cannot be used*	505
2. AI tools can be used in an assistive role*	506
3. AI has an integral role.....	506
UCL guidance on acknowledging use of AI and referencing AI.....	506
66. AI Guidance for Staff and Students.....	507
Generative Artificial Intelligence guidance for students	507
University position.....	507
Expectation of own original work	507
Important note.....	508
Current limitations of generative AI	508

It is important that you	508
Important note.....	509
Citing and acknowledging the use of AI.....	509
Further guidance	510
67. Generative AI in education.....	510
Generative AI.....	510
Practical information for lecturers.....	511
Disclaimer	511
Generative AI in our education	511
Key take-aways.....	511
What does this mean for your courses?.....	512
Misuse	512
Misuse of generative AI.....	512
When to suspect the use of generative AI tools	513
Using Turnitin to check for the use of generative AI tools.....	513
How does it work?.....	513
Step-by-step AI assessment guideline	514
1. Determine the risk in your assessment.....	514
2. Use generative AI to complete your assignment	514
3. Evaluate the output in relation to your assignment	515
Indications.....	515
4. Consider assessment criteria in relation to the output.....	515
Indications.....	515
Assessing writing skills.....	515
Communicating an AI-course policy.....	516
1. Determine the allowed use.....	516
2. Explain why you decided on the policy.....	516
3. Decide on the need for referencing, reporting, and/or reflection.....	516
68. Responsible use of generative Artificial Intelligence.....	517
Responsible use of generative Artificial Intelligence	517
What is generative AI?	517
Basic principles	517
Specific principles and guidelines per target group.....	518
The necessity of correct referencing.....	518
Tips and considerations.....	519
69. AI in education, resources for teaching faculty.....	519
AI in education.....	519
70. AI tools and your studies	520

About ChatGPT	520
Why does the UvA not use ChatGPT for teaching purposes?	520
AI tools may not be used to write assignments you'll be submitting.....	520
Assignments that you didn't write yourself may be deemed fraudulent	520
Risks when using AI tools	521
71. Guidelines to use artificial intelligence at UiO	521
How to use AI as a teacher	522
Understanding AI and the possibilities the tools provide	522
Assessment activities.....	525
Learning activities with AI?	528
How to Use AI as a Student.....	532
Learning with GPT UiO	532
Understand the Technology	532
Basics and Useful Resources.....	533
Explore and Use.....	534
Prompting.....	534
Proper, critical and honest use.....	535
Principles for the Use of Artificial Intelligence (AI) in Education.....	536
1. Safe and Responsible Use of AI.....	536
2. Choose AI Tools that Support Students' Learning Outcomes.....	536
3. Equal Access	537
4. Promote Critical Thinking	537
5. Speak with your students	537
6. Stay Updated	537
Ethics and Social Responsibility.....	538
Important questions you should reflect on	538
Can I input personal information into an AI tool?	538
Do I know who is behind the AI tool, and what interests are at play?	538
Can I input someone else's text or results into the AI tool?	538
Can I use AI tools to write my academic text and claim the output from the it as my own work?	539
Does Uio have rules and guidelines about cheating and dishonesty that I need to be aware of?	539
Can I trust the answers I get from artificial intelligence-based AI tools?	539
Legal guidelines for the use of Artificial Intelligence (AI)	540
Data Security	540
Copyright	542
72. Using AI to support learning	542
What are AI-based applications like?.....	542
University encourages the use of AI	543

Artificial intelligence guidelines in a nutshell	543
What are the guidelines based on?	544
73. Research in Artificial intelligence	545
74. Guidelines on using AI-powered chatbots in education and research.....	546
Guidelines on using AI-powered chatbots in education and research.....	546
AI chatbots and examination.....	547
Suspicion of cheating.....	548
Use of AI chatbots by teachers and students during courses	548
Use of AI chatbots by researchers in research and for research applications	549
How will SU continue to work on the issue	549
Help and support on AI chatbot issues	550
75. DTU opens up for the use of artificial intelligence in teaching	550
Students welcome AI	550
Guide for AI	551
Teaching needs to be personalized.....	551
76. AI chatbots in unsupervised assessment.....	552
About AI tools.....	552
How to assess assignments and projects	555
77. Artificial Intelligence in education.....	558
78. Guidelines for the use of generative AI	560
Guidelines for the use of generative AI	560
What can generative AI be used for?	560
AI for science communication.....	560
Risks	560
Privacy.....	561
Copyright	561
Lack of responsibility and accountability	561
Incorrect or discriminatory information.....	561
Dependency on technology.....	561
Safety	561
Guidelines for the use of generative AI	561
a. Privacy en data protection.....	562
b. Intellectual property	562
c. Accountability and Transparency	562
79. Guidelines for the Use of AI Tools.....	563

1. Artificial Intelligence for Teaching & Learning

The availability and uptake of generative Artificial Intelligence (AI) tools such as ChatGPT is impacting on the ways in which staff and students in universities teach, learn and assess. There are both risks and opportunities for an educational future shaped by the availability of AI tools, and both staff and students need to be aware of these in the context of the field of study and teaching of their discipline. The technologies themselves are rapidly evolving but there is an emerging consensus that students and staff need to develop ways of ethically using these technologies, be aware of the constraints and limitations but also the potential for innovation and enhancement in teaching and learning.

In response to teaching staff questions and as a way of keeping abreast of developments and issues, we have developed a series of guides for both staff and students.

AI Resource Guides

These guides are regularly updated to keep up with the fast-changing AI landscape. Written during May-July 2023, they aim to provide current information and will be continuously refreshed to include significant developments.

1.1. Staff Guide: Assessment and academic integrity in the age of AI

Discover our staff guide addressing the challenges surrounding assessments in light of the widespread accessibility of generative artificial intelligence tools. This guide delves into various issues and offers practical strategies, approaches, and recommended tools to safeguard academic integrity. Stay informed and learn how to effectively mitigate the potential threats posed by AI tools in the assessment process.

AI tools such as ChatGPT have sparked vigorous debates about assessments at universities. There are concerns around plagiarism and cheating as well as questions when it is appropriate for students to use AI. There have been calls for changes in the ways assessments are conceived, while others propose a return to paper-based only assessment strategies. Although AI tools have been on the

horizon, access to ChatGPT compelled educational institutions to develop responses. How we respond will vary, so this guide presents a range of suggestions.

[ChatGPT](#), which stands for Chat Generative Pre-Trained Transformer uses neural transformer networks to generate text. It and other AI tools such as Bard and Bing Chat are built on a category of AI tools known as ‘Large Language Models’ (LLM). These can perform various natural language processing tasks, such as, generate and classify text, provide answers to questions in a conversational style and translate texts from one language to another. These types of tools are commonly known as ‘generative AI’ which distinguishes them from other types of AI, already widely used in education (eg. writing aids such as Grammarly or Quillbot). The use of generative AI, such as ChatGPT requires the user to insert a language prompt. ChatGPT then splits out the words and makes a prediction of the best answer to the prompt based on the information it was trained on.

AI tools like these might help to minimize some repetitive or administrative tasks. However, these tools can still pose risks of giving inaccurate information or producing a low-quality output. This despite the very confident way results are presented.

Students have responded to these easy-to-access, powerful capabilities by using the tools for assessments, raising concerns about academic integrity. While the AI landscape and its effects on higher education are still being unpacked, there are attempts to circumvent the use of AI, such as by returning to more personal, handwritten and invigilated in-person exams.

There have also been calls for adapting some assessments so that they discourage the use of AI. For example, if assessments involve a relatively simple fact recall that generative AI can respond to, there may be enhanced ways to assess which can reduce tools like ChatGPT’s capacity to generate plausible answers. For instance, could the assessment also test students’ abilities to compare, use and analyse this knowledge? Or could an assessment include an interactive short oral presentation that can be scalable and effective? These strategies may not always be appropriate or practical.

Communicating with your students

AI literacies are becoming an essential skill. This involves developing an understanding of how the benefits must be balanced with the drawbacks. There is a need to equip our students with the necessary skills to navigate AI for their discipline.

At the beginning of your course, communicate to your students what you consider to be appropriate uses of generative AI tools for their assessments. There will be a spectrum across the university. This may include a requirement for a declaration that either such AI tools have not been used or that the use of these tools be [cited appropriately](#). If the use of AI is encouraged, you may want to specify tools or how/when students may use it.

Consider conveying the following to your students:

Specify whether generative AI tools can be used at the start of the course. It will be useful to have a discussion with your students about this to allow for questions. Lance Eaton's '[Syllabi Policies for Generative AI](#)' is a collection of statements used by different courses globally.

Specify how the AI tool should be referenced (e.g., APA style) and if further declarations should be made (i.e., including the prompts used).

Include conditions in the course plagiarism declaration/honour pledge or in the assignment submission instructions (see below).

Provide further guidance to students on using AI ([CILT student guide](#)).

Explain the consequences of academic dishonesty and inappropriate use of AI . (If you plan to make use of AI detection tools in the course, explain how you will go about verifying - see below). The following adapted standard declaration can be used for your course with the additional clause on third party and software to generate assessments. You may use or alter this statement to suit the course's needs. It is not recommended to ban the outright use of artificial intelligence software as many existing tools (e.g., grammar and style tools) are AI-based. However, one can restrict the generation or creation of an assignment using AI software.

Student plagiarism declaration

Example of adapting a plagiarism declaration to include using artificial intelligence software:

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

I have used the convention for citation and referencing. Each contribution to, and quotation in, this essay/report/project/ from the work(s) of other people

has been attributed, and has been cited and referenced. Any section taken from an internet source has been referenced to that source.

This essay/report/project/..... is my own work, and is in my own words (except where I have attributed it to others).

I have not paid a third party to complete my work on my behalf. My use of artificial intelligence software has been limited to (specify precisely how you used AI to assist with this assignment).

I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

I acknowledge that copying someone else's assignment or essay, or part of it, is wrong, and declare that this is my own work.

Approaches used in other university plagiarism declarations may also be considered. The [Liverpool University](#) declaration specifies 'no commissioned production' that includes ChatGPT. [Queens University, Belfast](#) has a simple statement "I certify that the submission is my own work, all sources are correctly attributed, and the contribution of any AI technologies is fully acknowledged".

Academic integrity strategies

Current generative AI can be used to 'cheat' on multiple choice quizzes, essays or coding assessments, and will undoubtedly continue to evolve. There are different ways to approach this challenge, as outlined by Michael Webb from JISC's National Centre for AI in [A Generative AI Primer](#). Where possible, the 'embrace and adapt' approach seems more likely to be effective and successful going forward.

How students use generative AI tools for support

Understanding more closely how students use generative AI tools for assessments is important. Popular uses of generative AI include summarising and paraphrasing long readings, generating ideas for assessment prompts, writing code, spelling and grammar checks (like Word and Grammarly), and generating practice questions for assessments. This may be analogous to how

Wikipedia gives an introductory overview of a topic, which may be less of a concern or even encouraged.

AI detection tools

There is a concern that students are increasingly using AI tools to complete their assessments. In response, AI detection tools are being developed to flag parts of written assignments that could possibly have been AI generated, say for example, using ChatGPT. Unfortunately, the accuracy of these tools is generally poor. This accuracy is certainly not comparable to a plagiarism checker such as Turnitin.

AI text detectors do not work in the same way as plagiarism checkers. A plagiarism checker such as Turnitin checks human generated text against other human generated text. AI text detectors estimate or guess the *probability* of text being written by AI. Some detectors were trained on datasets of human-written text and AI-written text to predict the probability that any given text was written by AI. Others work by looking for indicators within the text, such as linguistic patterns, repeated words or structural patterns.

When using plagiarism detection or originality tools to identify AI generated text:

There is a risk of falsely accusing students of plagiarism because detection errors are high.

Assume any detector will be unreliable given the rapidly evolving changes in AI.

Sharing student data by uploading it into a detector that is managed by a third-party company with unknown privacy and data usage policies is ethically and legally risky.

In an empirical review of AI text generation detectors, [Sadasivan et al. \(2023\)](#) found “that several AI-text detectors are not reliable in practical scenarios”. In a research study [Ibrahim et al. \(2023\)](#) state that “current AI-text classifiers cannot reliably detect ChatGPT’s use in school work”. The widespread use of other types of AI, such as Grammarly or Quillbot, further undermines accurate detection. The researchers found it was relatively easy to evade detection by running ”the ChatGPT- generated text through Quillbot— a popular paraphrasing/ rewriting tool utilized by students worldwide” resulting in rendering “these algorithms futile, failing to detect 95% of ChatGPT answers.”

There are freely available tools that given AI generated text will reduce the chances of this being detected as AI generated. These are known as AI content detection bypassers. These tools claim to be able to “humanise” your AI generated text, often through paraphrasing.

Given these issues, there is an emerging consensus within the academic community that detectors are not reliable and may cause harm to particular groups of students. AI detectors have been found to be more likely to [label text written by “non-native English” speakers as AI-written](#). In the study by Stanford researchers, further revealed a trend “where more literary language was classified as more ‘human’: enhancement of word choice in non-native English writing samples reduced misclassification, while simplifying native writing samples increased it, suggesting that GPT detectors are inadvertently penalizing individuals with limited linguistic proficiency.” (Liang et al., 2023).

Several AI detection tools are available, including Turnitin, Copyleaks, Content at Scale and Originality.AI, however none of these has proved totally reliable. Initial claims from the vendors suggested high accuracy but real world testing and empirical is revealing these purported accuracy rates are unreliable and need further scrutiny:

Turnitin has backtracked on the initial claims [after wider usage and testing](#). A number of universities have switched off the Turnitin detector including Wits and [Vanderbilt University](#).

OpenAI has [shut down its own AI detector](#) due to low accuracy rates.

A sample of the more commonly used detectors’ webpages or terms of reference indicate they should *not* be used as a primary mode to determine whether a text is AI-generated (source [AI detectors slide deck](#) from Torrey Trust)

Recommendations on AI detection tools

AI detectors should not be relied upon for accurately ascertaining or giving a definitive result as to whether text is AI or human generated.

AI detection reports cannot be relied on as stand-alone evidence of whether a student submission has been AI generated in a disciplinary context.

Staff should use AI detectors with care bearing in mind:

the possibility of false positives,

the relative ease in which detection can be evaded,

AI detectors may disproportionately flag text written by non first language speakers as AI generated.

Risks and concerns of using AI

There are wider risks and concerns to be considered if you plan to use AI in your classroom that may not be immediately apparent.

Relevance of data set: Generative AI may not have the latest information or developments in a field or topic. For example, ChatGPT version 3.5 was trained on data up until 2021. When proposing using AI tools, you should double-check its outputs and consult sources to ensure that the information provided is appropriate and accurate.

Inaccuracies in what is generated: While AI outputs may come across as authoritative and convincing, its responses are based on next word predictions. The software has no real understanding. This leads to ‘hallucinations’ and allows it to produce disinformation (see [why ChatGPT produces inaccuracies](#)). It is important to evaluate this risk in the context of how you intend using generative AI.

Bias in data fed: AI follows the ‘garbage-in-garbage out’ principle whereby if the data it has been fed is biased, then its responses will reflect that bias. There are a number of biases that are present in AI data, such as, historical, representative, algorithmic, ranking, behavioural and social biases (read more about [bias and fairness in AI systems](#)). You need to be aware of the potential bias in AI outputs and critically evaluate the information it produces.

Dependence/Overreliance: Assignments serve the purpose of helping students learn and practice skills that we will need in society. While AI can help automate some tasks, it is important that students do not become overly reliant on it as this may hinder their ability to craft and think critically. Research suggests that being too reliant on AI can lead to loss of some important cognitive skills and prime us to think in a certain way. (See [ChatGPT can homogenise our lives](#)).

Designing assessments for higher order thinking

Courses involve different levels of knowledge. Generative AI is far less capable at higher levels. Table 1 shows examples of opportunities to adapt assessments to target these higher forms of learning. Additional tips include reviewing the grading mechanism and rubrics and reducing the emphasis on recall.

Table 1: Assessment difficulty spectrum for AI (Adapted from Monash, 2023)

Difficulty level	Assessment type	Explanation of use
Easy	MCQ quizzes and questions involving recall	Generative AI can easily provide output to factoid or low-level questions, especially on widely taught topics.
Easy	Generic short written assignments	Generative AI can produce convincing essays and poems. For example, “Write a 1000-word essay on the history of Nelson Mandela”.
Medium	Scaffolded submissions	Creating a scaffolded assignment allows students to build on their previous work and feedback. Include pre-writing and drafting in the assignment process. Verifiable sources and citations should be required
Medium	Personalised or context-based assessment	A simple essay can be made more challenging by introducing personalisation. Encourage personalisation where students are asked to draw on personal experiences. Ask students to write to a particular audience whose knowledge and values must be considered amplifying their student voice. Ask questions that would require them to give a response that draws from concepts that were done in class, in a lab, field trip or real-life experiences in their contexts.
Hard	Projects	Projects involving real-world applications give students the opportunity for meaningful learning experiences.
Hard	Oral tests / exam with Q&A / panels or discussions	Synchronous oral assessment allows checking a student’s understanding of their submitted work and gives the opportunity to interrogate their submission with follow up questions and discussion. While time intensive for larger classes, some have found ways to use these strategies.

Using prompts for preparation

Generative AI has also been explored for routine assessment tasks and anticipating what students might find. There are, of course, limitations and concerns. Investigating these uses can help to understand how the AI tool works, what routine tasks can be automated, and how your students may use it.

Testing existing assessments

Test your assignment briefs in an AI generative tool to see if it can be easily completed. Review the results and fine tune your prompts as a student would. Look closely at the output as while it may seem reasonable it may lack the finer details required from assignment briefs.

Create quiz questions

Generative AI tools, given a lecture transcript or notes, can then be asked to generate multiple choice questions with answer options. It is usually good at generating the quiz format, the questions and options will require substantial editing and checking before being used in a course.

Prompt: "Read the following transcript from a lecture video and write 10 multiple choice quiz questions with 4 distractors each and feedback. Indicate the correct answer. This quiz should be pitched at a university level."

Suggestions for improving prompts include avoiding making them overly complicated as the output will become confused. One may also suggest to ChatGPT or similar tool to ‘take the task step by step’ and to ‘rethink’ earlier responses if the generated output is not satisfactory.

Create rubrics

ChatGPT, Bing Chat or Bard can be asked to generate a marking rubric. The prompt could include details such as the total marks and details of the task being assessed. This short video shows how to [Create a Marking Rubric With AI \(ChatGPT\)](#)

Prompt: “Generate a rubric out of 12 for a student assignment that requires them to create an app on Fliplet. These are university students studying in an education discipline.”

Generate writing prompts

ChatGPT or similar tools can generate possible writing prompts or sample data for a student assignment. Make sure to include the topic and theme in your prompt.

Prompt: “Generate five essay topics on using AI for teaching school mathematics. The essay should be 1000 words in length and use an example of how AI is used in schools

1.1.1. AI-proof checklist for your assessments

An AI misuse or proof checklist (Table 2) may assist when planning new assessments. No assignment can be made completely “AI-proof”, but clear communications and reducing the likelihood that students could use AI tools inappropriately can be planned ahead. This is not to underestimate the time and potential resource implications of making changes to AI-proof assignments.

Table 2: AI proofing checklist (Adapted from Turnitin, 2023)

Criteria
Does the assignment brief make explicit your institution and course’s academic integrity policy, especially regarding the use of AI text tools?
Does the assignment brief communicate the acceptable and unacceptable limits of using generative AI tools for the student response?
Does the assignment brief require critical thinking or reasoning? Does the assignment encourage/require student voice?
Does the assignment brief require the student to incorporate personal stories and/ or authentic situations?
Does the assignment brief require a list of verifiable sources and/or citations? Are students asked to include a reflection or rationale for their approach to the assignment solutions?
Have you instituted checkpoints to review outlines or drafts throughout the course, rather than focusing on a final submission only?
Have you included time for peer reviews and/or discussions on learning activities throughout the course or assignment?
Have you run the assignment brief through an AI generative tool?

1.2. Staff Guide: Teaching and learning with AI tools

Explore our comprehensive guide on the uses of generative AI tools like ChatGPT in teaching and learning. Discover potential applications, examine ethical concerns, and access additional resources.

Introduction

ChatGPT has grabbed the headlines since its public release in November 2022. ChatGPT (<https://chat.openai.com/>), which stands for Chat Generative Pre-trained Transformer, has a basic version available to anyone who signs up for a free account while a more functional version (eg. ChatGPT Plus) is being released as a subscription service. It can perform various natural language processing tasks, such as, generating and classifying text, providing answers to questions in a conversational style and translating texts from one language to another.

The AI definition has developed over time, but broadly AI is a range of technologies that perform cognitive tasks, through machine learning, natural language processing, data mining, neural networks or an algorithm (Zawacki-Richter et al., 2019).

Some claim ChatGPT will become the '[greatest cheating machine ever](#)', while others suggest it will open [new possibilities](#) to augment our skills by helping to automate mundane and routine tasks. More AI tools are being developed, for example, Bing Chat and Google Bard which also operate using LLMs, like ChatGPT but access information from the internet, and focus more on operating as search tools.

The explosion of interest reflects how the applications and uses of AI are rapidly developing – including being used by students. As educators, we need to be keeping up to ensure that our teaching and learning is providing students with the skills and knowledge they need to navigate the world of AI.

Setting expectations

Outlining what faculties, programmes or departments expect from students and staff around the use of AI tools in teaching and learning, especially before any assessments are undertaken is essential.

Instead of trying to stop students using AI, lay out expectations about the need to make visible and acknowledge the use of AI tools. Engagement with staff and students to promote debate and knowledge sharing is part of the development of new literacies for education and work.

Some higher education institutions have ‘banned’ the use of AI and students using AI are considered to have been academically dishonest ([University World News](#)), with severe penalties. Using AI detection tools (for example, Turnitin) has been adopted as a screening mechanism in some contexts. But detection tools are unreliable (see [discussion from Monash University](#) and CILT tests), and some universities explicitly refrain from using them because they risk students being wrongfully accused of cheating, or of evading detection. While academic integrity is a concern, taking a punitive approach is likely to drive practices underground rather than build key skills for the future.

AI is here to stay, and higher education has to adapt to incorporate the emerging uses of the new technologies. However, we need to be mindful of the ethical issues associated with using AI in education.

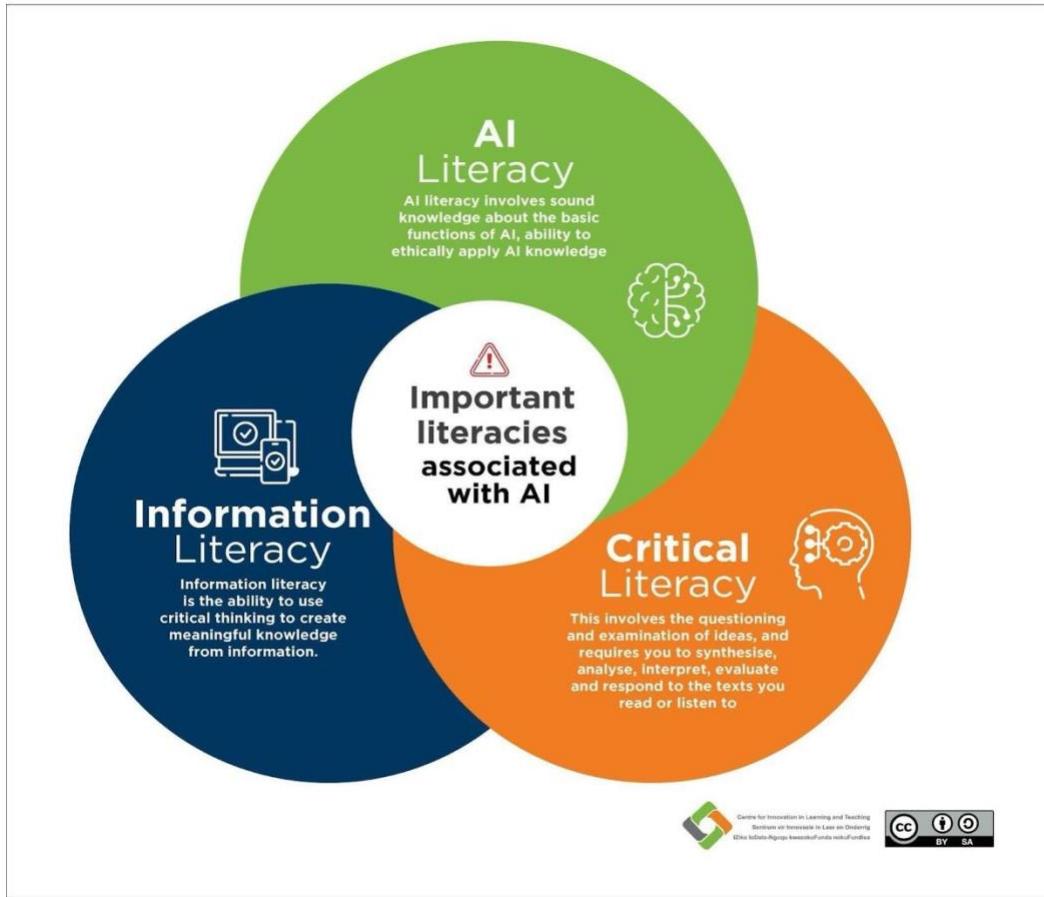
Ethics of Using ChatGPT (or similar tools) in education

- **Academic integrity** – There should be an emphasis on the distinction between using AI as an enabler or assistant and using AI as an authoring tool which impacts on academic integrity. Unless made explicit, it can be tempting for students to pass off AI generated work as their own, since it is not ‘copied’ as in the traditional form of plagiarism.
- **Privacy risks** – ChatGPT’s Privacy Policy clearly states that “By using our Service, you understand and acknowledge that your Personal Information will be processed and stored in our facilities and servers in the United States ...” noting that the US privacy environment/legislation according to general consensus understanding is deemed to not be equivalent or better when assessing it against our own legislation. Refer Section 72 of POPIA. The fundamental take away is no personal information should be submitted to ChatGPT.

- **Inaccuracies** – AI outputs are known to include inaccurate information, limited by the training dataset exposure and the AI’s capacity to ‘hallucinate’. This means that AI tools like ChatGPT have a high risk of spreading misinformation, given it has no capacity to distinguish right from wrong, correct from incorrect.
- **Built in-biases** – AI trained off data sets which contain implicit and explicit biases (overlooking local knowledge, lack of cultural diversity, dominance of Western hegemonic knowledge; baked in racism, sexism, and other undesirable values)
- **Exploitation of labour** – there have been allegations of exploitative labour practices in the training of AI
- **Accessibility** – there is a risk that those with poorer access to connectivity, devices, data and literacies will get unequal access to the opportunities being provided by AI. More powerful AI capabilities will be marketized and only available to those with resources (which has already happened with ChatGPT Plus and DALL-E, the visual generator). In a context with existing inequality, planning the use of tools such as these for teaching and learning should be cognizant of access.

Important literacies associated with AI

There are three core literacies (AI, Critical and Information literacy) required for all stakeholders to navigate AI effectively and ought to be considered as part of graduate attributes.



AI Literacy

AI literacy involves sound knowledge about the basic functions of AI, ability to ethically apply AI knowledge, concepts and applications in different scenarios and ability to critically evaluate AI technologies, communicate and collaborate effectively with AI (Ng et al., 2021). AI literacy includes understanding what AI can be used for, what it does not do well and how to acknowledge its use (see [Monash University library advice](#) and [APA citation style](#) for ChatGPT as an example).

Users need to understand the basic operation of the models and be familiar with the inherent limits of the systems they are making use of – for example, the lack of ‘common sense’ or real understanding which allows for unrealistic or inaccurate answers. An infographic developed by UNESCO provides basic guidelines about [when it is safe to use ChatGPT](#), which could also be applied to other AI tools.

A new skill required is learning “prompt engineering”, which means how best to formulate the questions that provide the parameters and guidelines for how the AI should respond.

Here are some guidelines for writing requests for generative AI summarised from an extended article on using [ChatGPT for Higher Education and Professional Development](#) (Atlas, 2023; pp.41-48):

- Choose your words carefully
- Begin by defining the purpose and focus
- Be specific and concise
- Provide context
- Ask for more; or stop and redirect

Critical literacy

This involves the questioning and examination of ideas, and requires you to synthesise, analyse, interpret, evaluate and respond to the texts you read or listen to (see the [University of Melbourne](#) for some ideas for interrogating text). Bearing in mind that the Large Language Models (LLMs) are based on predictions and only have access to the dataset of information they are trained on, and that AI generated information is not always accurate and authentic, students need to be critically literate. With the responses generated from the Large Language Models, which are usually well formulated in terms of grammar, syntax and natural language expressions, making them sound authoritative and plausible, the value of learning the skills for critical literacy, which is a traditional learning outcome of higher education, is even more important.

This may mean creating orientation modules, courses or components of courses that build students' capacity to effectively understand and apply AI tools both in their educational projects and their professional domain areas of expertise. Designing learning activities which challenge students to interact with AI to get a sense of its capabilities, and to evaluate its value and outputs can be incorporated into all courses and programmes. Treating AI as another tool in teaching and learning which students are expected to gain skills and competence in shifts it out of the underground world of cheating or short cuts to becoming a resource for achieving tasks.

Information literacy

Given the potential and capacity of AI in education, it becomes imperative that rapid upskilling all incoming students in information literacy - "the ability to use critical thinking to create meaningful knowledge from information" (Claremont Colleges Library Information Literacy Steering Group) - is a baseline part of their induction into university. That is, students need to be able to identify what information is required to complete their assignments or assessments, and then evaluate this information and communicate it in an ethical and legal manner (see how this group frames the Critical Information Literacy Habits of Mind).

Early assessment and training on information literacy for effective learning is critical. Ensure that all first-year students attend visits to the UCT Libraries for such assistance.

AI and academic integrity

The use of AI tools by students for assessments is being seen as a challenge in higher education globally. The current AI capabilities can be used to 'cheat' at simple multiple choice, basic essay or coding assessments, and will undoubtedly continue to improve. As educators, there are several ways to approach this challenge, as outlined by Michael Webb from JISC's National Centre for AI in [A Generative AI Primer](#).

 Strategy	 Approach	 Challenges
Avoid	Revert to in-person assessments where the use of AI isn't possible	This moves away from authentic assessment and creates many logistical challenges
Outrun	Devise an assessment that AI can't do	AI is advancing rapidly and given the time between the assessment being set and it being taken, AI might well be able to do the assignment when it is taken
Embrace and adapt	Embrace the use of AI, discuss the appropriate use of AI with students, and actively encourage its use to create authentic assessments	Balancing authentic assessment and the use of generative AI with academic integrity is a challenge

Where possible, the ‘embrace and adapt’ approach seems more likely to be effective and successful. In the companion CILT guide: [Staff Guide to Assessment and Academic Integrity in the age of AI](#) we outline in more detail ways in which you can design and think about assessments that embrace and adapt to AI tools.

Roles and uses to explore for AI tools in teaching and learning

While we must be responsible in how we use the emerging capabilities of technology, there are also interesting new opportunities to be explored, depending on your context and your students. The list of use cases below is not exhaustive but provides a glimpse of the enormous potential for deployment in teaching and learning.

Staff	Students
Content Creation	Adaptive Learning Materials

Using AI to generate educational materials such as lesson plans, quizzes, and interactive content	Students accessing adaptive learning materials that adjust difficulty levels based on their progress
Intelligent Tutoring Systems Deploying AI-powered tutoring systems that adapt to individual student needs	Virtual Mentors Accessing virtual mentors, powered by AI, to provide guidance and support throughout their learning journey
Automated Administrative Tasks Utilising AI for automating administrative tasks such as scheduling, and record-keeping	Language Learning Support Using AI-based language learning tools that simulate conversations, provide instant feedback, and generate language exercises
Adaptive Learning Paths Creating personalised learning paths for students based on their individual progress and learning styles	Personalised Learning Paths Students receiving tailored learning paths based on their individual strengths and weaknesses
Virtual Reality Simulations Integrating AI with virtual reality technology to create immersive educational simulations	Personalised Assessments Students taking assessments that dynamically adjust difficulty levels based on their performance and provide personalized feedback
Automated Quiz Grading Utilising AI algorithms to automatically grade objective assignments	Study and Homework Assistance Accessing AI-powered study and homework assistance tools that provide explanations, suggestions, and resources
Learning Analytics Applying AI to analyse learning patterns and provide actionable insights to optimise instruction	Skill Development and Practice Using AI-based platforms to develop and practice specific skills, such as coding or language proficiency

Rubric categories Generate a marking rubric for an open-ended topic	Personalised Study Schedules Students receiving AI-generated study schedules tailored to their preferences and optimal learning times
Natural Language Processing Utilising AI's natural language processing capabilities for tasks like automated language translation, question-answering systems, and text analysis	Concept Understanding and Reinforcement AI-powered tools that help students understand and reinforce complex concepts through interactive explanations, visualisations, and practice exercises

The [UNESCO Quick Guide](#) has also outlined 10 educational ‘roles’ for AI tools ranging from ‘Socratic opponent’ to ‘study buddy’ and ‘dynamic assessor’.

If you make use of AI for teaching, please share your experiences [in this form](#), so we can create new appropriate cases from our own context.

What AI tools are there for education?

While ChatGPT has won the media headlines, there are many ways in which AI is being built into a variety of tools and platforms to use in education and industry. Here are a few examples for you to explore (some may require a payment while others are free).

- Interactive ([ChatGPT](#), [YouChat](#), [ChatSonic](#))
- Writing/text ([QuillBot](#), [Sudowrite](#), [WriteSonic](#), [Jasper](#), [You](#), [Moonbeam](#))
- Visual AI ([DALL E2](#), [Midjourney](#))
- Research and publishing ([Perplexity AI](#), [Researcher Life](#), [Elicit](#), [Sci Space](#))

The [Futuretools](#) website offers an extensive list of AI tools for various activities (you can filter for free tools).

1.3. Student Guide: Using ChatGPT and other AI tools in education

Explore our comprehensive guide on the uses of generative AI tools like ChatGPT in teaching and learning. Discover potential applications, examine ethical concerns, and access additional resources.

The purpose of this guide is to take you through some of the current debates about the risks and benefits of using artificial intelligence and more particularly, ChatGPT for learning. To effectively make use of the ChatGPT tool, it is important to know how it generates its outputs.

What is AI?

The Artificial Intelligence (AI) definition has developed over time, from “the science and engineering of making intelligent machines” in 1956 (McCarthy, 2007, p.2) to broader ones describing AI as a range of technologies that perform intellectual learning tasks, such as machine learning, natural language processing, data mining, neural networks or an algorithm (Zawacki-Richter et al., 2019). AI has become an integral part of various applications and settings. As AI evolves, a powerful branch called generative AI is emerging. This type of AI can create new content, like text, images, and code. [ChatGPT](#), [CoPilot](#) and [Google Bard](#) are examples of generative AI.

What's the big deal with generative AI such as ChatGPT?

ChatGPT (<https://chat.openai.com/>), which stands for Chat Generative Pre-trained Transformer, has gained a lot of attention since its release in November 2022. Currently, anyone can sign up for a free account using version 3.5, although improved versions (eg. ChatGPT Plus) are a paid-for service.

It is a software in the form of a chatbot that was developed by [OpenAI](#). It is built on a category of AI tools known as ‘[Large Language Models](#)’ (LLM), which can perform various natural language processing tasks, such as generating and classifying text, providing answers to questions in a conversational style and translating texts from one language to another.

Some claim ChatGPT will become the '[greatest cheating machine ever](#)', while others argue that it will open [new possibilities](#) to enhance or improve our skills as the tool that can help automate mundane and routine tasks.

More AI tools are being developed, which also operate using LLMs, like ChatGPT but access information from the internet, and focus more on operating as search tools.

How does ChatGPT generate its outputs?

ChatGPT is a software that generates its outputs using natural language processing algorithms. When a user inputs a prompt, it analyses the prompt using its trained knowledge and generates a response. ChatGPT's trained knowledge is based on a vast amount of text data that it has been fed from the Internet. Its outputs are thus only as good as the data it has been trained on. If the training data contains misinformation and biases, then its outputs will reflect those issues.

As a software which produces its response based on algorithms, it lacks understanding of real-world experiences, social and cultural distinctions. The responses it produces thus may not be contextually relevant or lack real understanding. These mechanisms that govern how ChatGPT produces its responses lead to some risks in using ChatGPT or similar AI tools.

Risks and Ethics of using ChatGPT (or similar tools) in education

While ChatGPT can help you with your learning or preparing for an assignment, its outputs may be outdated or contain errors and biases. Here are some risks associated with using ChatGPT:

Inaccuracies in data fed: While ChatGPT's outputs may come across as reliable and convincing, its responses are based on next word predictions. The software has no real understanding, and no capacity to distinguish right from wrong, or correct from incorrect. ChatGPT-3.5, as an AI has the capacity to 'hallucinate' and produce misinformation. It is important to evaluate and verify ChatGPT responses and not take the information generated as given. Read more about [why ChatGPT produces inaccuracies](#).

Built-in biases: AI is trained off data sets which contain implicit/indirect and explicit/direct biases, including overlooking local knowledge, lack of cultural diversity, dominance of Western influential knowledge; baked in racism, sexism, and other undesirable values (read more about

[bias and fairness in AI systems](#)). ChatGPT follows the ‘garbage-in-garbage out’ principle, meaning if the data it has been fed is biased, then its responses will reflect that bias. You need to be aware of the potential bias in ChatGPT’s outputs and critically evaluate the information it produces. There have also been allegations of unfair/ exploitative labour practices in the training of AI.

Relevance of data fed: ChatGPT-3.5 is trained on data that is publicly available up until January 2022. It, thus, may not have the latest information or developments in a field or topic. When using ChatGPT, you should double-check its outputs and consult sources to ensure that the information provided is up-to-date and relevant.

Dependence/Overreliance: Assignments serve the purpose of helping us learn and practice skills that we will need in society. While ChatGPT can help automate some tasks, it is important that you do not become overly reliant on it so that it hinders your ability to craft and think critically for yourself. Research suggests that being too reliant on ChatGPT can lead to loss of some important cognitive skills and prime us to think in a certain way. You can read more about how [ChatGPT can homogenise our lives](#).

Now that you know how ChatGPT generates its outputs and the risks and ethical considerations associated with it, let’s look at how you can effectively use the tool to help you learn better.

How ChatGPT can help you in your learning processes?

AI tools are here to stay and can be used to aid your learning but should not be used to complete your assignments. There are several ways ChatGPT can be used effectively in your learning, such as the examples below:

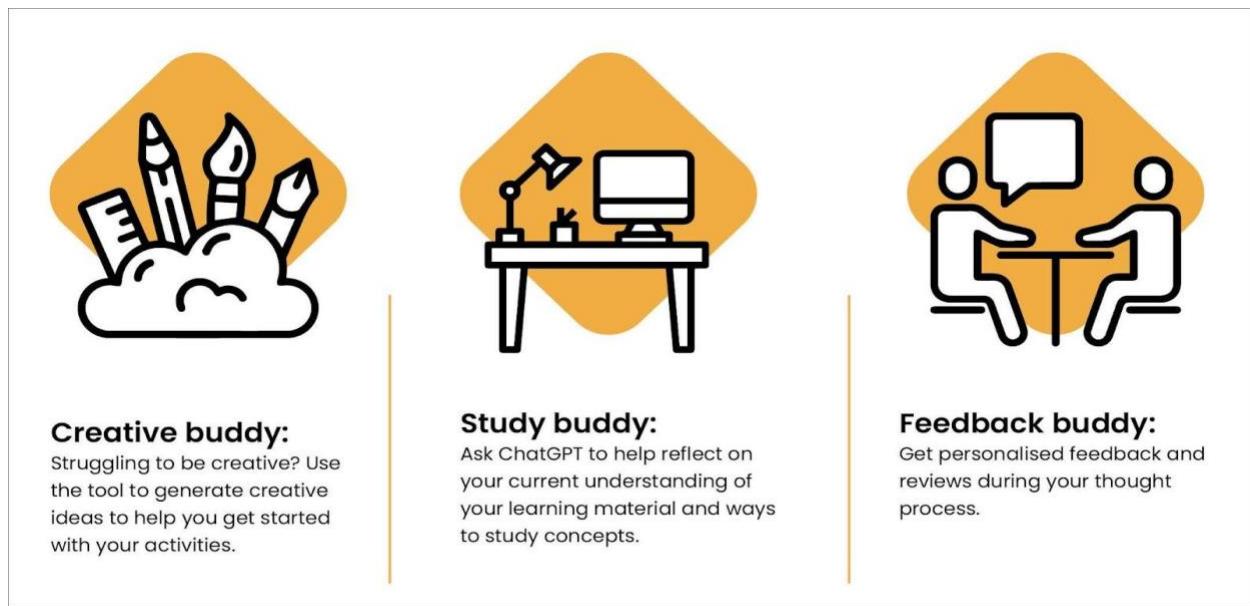


Figure 1: Less-risky ways of using ChatGPT for effective learning (adapted from the 2023 UNESCO [Quick Start Guide](#): ChatGPT and Artificial Intelligence in higher education)

Here are other AI tools to explore:

Generative AI ([Google Bard](#), [CoPilot](#))

Writing/text ([QuillBot](#), [Sudowrite](#), [WriteSonic](#), [Jasper](#), [You](#), [Moonbeam](#))

Visual AI ([DALL E2](#), [DALL E3](#), [Midjourney](#))

Media generation ([AI-generated art](#))

Research and publishing ([Perplexity AI](#), [Researcher Life](#), [Elicit](#), [Sci Space](#)),

However, it is important for you to know *what is permitted in your course* so you can make informed decisions. To begin with, check the university, department and course policies around academic integrity and plagiarism, and ask your lecturer to clarify about the use of AI for course assessments.

AI Tools and Academic Integrity

You need to make sure you do not use AI tools inappropriately and be accused of plagiarism. It may be tempting to pass off AI generated work as your own, since it is not ‘copying’ other students’ work but unless you declare it, generating assignments using ChatGPT means it is not your own work. Assessments are designed to test your understanding of a subject and your ability

to apply that knowledge and to prepare you for the workplace. If you use AI to simply write your assessment the value of the assessment would be diminished; you may not develop the required skills for your future studies or career, and you could be accused of plagiarism.

“Plagiarism is the misappropriation of others’ words, thoughts and ideas by presenting them as one’s own, and is treated very seriously in the academic world. Under no circumstances is it acceptable to present the work of others as your own.” [UCT Author-date Reference Guide: based on the Harvard referencing style](#) (2016)

The distinction between using AI as an enabler or assistant and using AI as an authoring tool has an impact on academic integrity – using AI tools for assistance can be done effectively without plagiarising. Not all lecturers and supervisors will have the same views, so check first, but the following are some examples of appropriate and questionable uses of generative AI.

D	Uses of generative AI at university likely to be acceptable	DON 'T	Uses of generative AI at university likely to be unacceptable
✓	Find out whether or how AI tools can be used for each assignment before starting	x	Use AI tools when they are specifically prohibited
✓	Attribute the use of AI (eg <u>APA style</u>)	x	Use AI tools without acknowledgement
✓	Use AI as a prompt for an outline (unless specifically prohibited)	x	Copy an AI output and pass it off as your own work
✓	Brainstorm ideas or request summaries of information	x	Trust AI outputs without doing a critical check for facts and sources
✓	Ask ChatGPT to rephrase a difficult concept into simpler language	x	Share any personal information or upload copyrighted materials
✓	Save your prompts and the outputs in case you are challenged about your use	x	Use AI tools when original content is being expected

First ask the lecturers and tutors for guidance about using AI tools for the course assessments.

Always validate outputs from ChatGPT or other AI tools – that is, applying [critical literacy](#) skills to evaluate content.

Make sure that the final product is your own work, and not just copied from an AI generator. You can use the generated text as a learning tool to formulate initial ideas or to suggest an essay structure, but the final submitted assessment must be your own work, creation, and analysis.

You must appropriately [acknowledge](#) where AI generative tools have been used in an assessment or any other part of your work, clearly indicate where you have used the AI generative tool and the extent to which it was used. Using AI-generated content without acknowledgement is a breach of academic integrity that may result in academic misconduct allegations and subsequent consequences.

The UCT Code of Conduct

[UCT RULES ON CONDUCT FOR STUDENTS](#)

Student Rules - Academic conduct

RCS2.3

A student may not submit the work of any other person in any examination, test or in respect of the completion and/or submission of any other form of academic assessment without full and proper attribution and acknowledgement - dishonest conduct.

NOTE: Guidance on forms of referencing is available from academic staff, the staff of the UCT Libraries and from the Writing Centre.

1.4. Student Guide: Developing effective prompts for generative AI tools

Prompting involves crafting effective strategies to engage with generative AI tools. Clearer instructions result in more accurate outputs. Ambiguous prompts are more likely to produce inaccurate outputs. This guide summarises productive prompting strategies and common practices.

Prompting involves crafting effective strategies to engage with generative AI tools. Clearer instructions result in more accurate outputs. Ambiguous prompts are more likely to produce inaccurate outputs. This guide summarises productive prompting strategies and common practices.

Understanding how generative AI works

The natural language style of generative AI tools invites us to “think about a prompt as a conversation” (White, 2023a; Cathey, 2023). However, these are NOT human conversations, even though they mimic human interactions. Generative AI uses large language models (LLMs) that are responding to your inputs by trying to predict a next likely word (see [A Generative AI Primer](#)). By crafting prompts one can control and direct what is generated so it becomes more useful.

The growing number of generative AI tools (for example, [ChatGPT](#), [Bard](#), [Bing Chat](#), [Claude](#)) all have some differences around how they operate and the data on which they were trained. One large distinguisher is the amount of prompt information they permit, some also allow you to upload documents or images to assist with the tasks. Trying more than one platform is useful to get a sense of their capabilities and weaknesses.

Whichever platform you select, the prompting strategies will be similar. We have included a short, ‘cheat sheet’ approach below, and a longer explanation of specific prompt patterns you can adopt.

Prompts ‘cheat sheet’

You can use the acronym CREATE, as a shorthand approach to writing prompts: Clarity, Relevant info, Examples, Avoid Ambiguity, Tinker, Evaluate ([Barrett, 2023](#))

Clarity	Clearly define the task or intent of the prompt, including specific information about the output.
Relevant info	Provide relevant details, including specific keywords and facts, the tone, audience, format, and structure.

Examples	Use examples in the prompt to provide context and direction for the output.
Avoid ambiguity	Focus on the key information and delete unnecessary details in the prompt.
Tinker	Test and refine the prompt through multiple iterations. Explore different input versions to discover the best results. Try different platforms.
Evaluate	Continuously evaluate the output and adjust the prompt as needed to improve the quality.

Giving a context is more likely to generate the results you anticipated. The following pattern strategies are more directive and can be used when prompting.

Useful prompt patterns strategies

Prompt engineering is an emerging discipline which refers to the development of prompts to efficiently use generative AI tools for a wide variety of applications. But anyone can draw on these ideas to improve the phrasing of queries to generate outputs. There are many emerging frameworks for prompt engineering, but since this guide is not concerned with the technical aspects, we simply use the term ‘prompting’ to refer to crafting the best approach to engage a generative AI tool.

The guide offers an introduction to some of the most widely used pattern-based strategies for AI tools (White et al., 2023; Corral, 2023; Alattas, 2023; Ismuguzel, 2023). Each of the strategies is briefly explained and a longer illustrative example of how this pattern can be used is linked to the Addendum.

The strategies are roughly grouped into three prompting pattern types:

Patterns based on inputs and outputs.

Role playing

Audience

Patterns based on refining and interrogating the questions.

Question refinement

Cognitive verification

Flipped interaction

Patterns based on specifying a template.

Patterns based on inputs and outputs

These are prompts that specify what kind of input the model should expect and what kind of output it should generate. For example, the role-playing or persona pattern requires you to ask the AI tool to assume the role of an expert, or a specified character. The audience pattern defines exactly for whom the output is intended. You could combine both patterns in a series of prompts.

Role-playing Pattern

This type of prompt instructs the AI tool to respond from the perspective of a specific persona or role. This may range from assuming the appropriate tone to having the AI tool take on a coaching or teaching role for you. In specifying a role, you can also formulate your expectation of the output. Before prompting, you may want to also provide examples of your writing style, or the output style you want.

Best for: delivering more tailored and accurate outputs with the correct tone

Here is an example of a role-playing prompt:

Your role as a coding tutor is to create personalized study plans to help first year university students learn how to code in the Python language. Your responsibilities will include understanding the goals, time commitment, and preferred learning resources of each student, and using that information to develop a comprehensive study plan with clear timelines and links to relevant resources. You should be able to adapt your teaching style to meet the individual needs of each student and provide ongoing support and guidance throughout the learning process. Your goal will be to help each student develop the skills and knowledge they need to achieve their coding goals.

Prompt format: Act as persona X. Provide information about topic Y. Nyakundi (2023) suggests using the 5 Ws framework in formulating role-based prompts.

Who – specify the role you need the AI tool to assume.

What – indicate the action you require the AI tool to perform.

When – specify the timeframe in which the AI tool must complete its task.

Where – specify the location or context of your prompt.

Why – articulate the motivation, reasoning, or goals governing the prompts.

Remember that the more specifications you give, the more tailored results you will get. Tip: The role does not have to be human; you can ask the AI to pretend to be an inanimate object like a database or a Linux machine.



Worked Example: [Applying the Role-playing](#)

[Pattern](#)

Audience Pattern

The Audience Pattern moves the focus to who is reading the outputs. When formulating this type of prompt, you should request information from the AI tool for a specific audience (Corrall, 2023, White et al., 2023).

Prompt format:

Explain X to me.

Assume that I am persona Y.

Best for: Tasks such as summarizing, translating, paraphrasing, or captioning.

Prompt format example:

Please help me write helpful information for university researchers. The writing style should be similar. I want you to understand this writing style so I will provide some examples. Remember my writing style as being PATTERN_STYLE. Afterwards when I ask you to write using PATTERN_STYLE, then use this style of writing. (afterwards provide example paragraphs)



Worked Example: [Applying the Audience](#)

[Pattern](#)

Patterns based on refining and interrogating the questions

Another approach is to focus on refining and interrogating the questions you give the AI tools and making use of the AI's capacity to enhance your prompts. Three related patterns are the question refinement pattern, the cognitive verification pattern and the flipped interaction pattern which asks the AI to come up with better formulated questions, asks the AI to fact check its answers or asks the AI to solicit more information from you.

Best for: Approaching complex tasks or ambiguous queries, providing explanations or justifications, or checking facts or logic.

Question Refinement Pattern

This type of prompt enables you to utilize the AI tool's capabilities to help in framing a question to arrive at an accurate answer (White et al., 2023) by suggesting a better version of the question to use instead.

Prompt format:

When I ask a question, suggest a better version of the question to use instead.

Optional: Ask me if I would like to use the better version instead.

Applying the Question Refinement Pattern



Worked Examples: [Exchange 1](#) showcases a basic version of this.

This method of prompting can also be further refined for specificity, [see Exchange 2](#).

Cognitive Verification Pattern

This type of prompt has the user providing the AI tool with rules on how to respond to secure a comprehensive answer. It involves subdividing the initial question into smaller ones that may aid in informing the response. The aim is to improve the accuracy of the AI tool's response to the main question by forcing it to incorporate potentially omitted pieces of information.

Prompt format example: (adapted from White et al., 2023: 5):

When I ask a question, follow these rules. Generate a few additional questions that would help you answer the question more accurately. Combine the answers to the individual questions to produce

the final answer to the question. Assume I know little about the topic we are discussing and please define any terms that are not general knowledge. When I have answered three of your questions, combine the answers to produce the final answer to my original question.

Step 1: Start by specifying ‘when you are asked a question, follow these rules’.

Step 2: Generate sub-questions that would help answer the question more accurately.

Ask the AI tool to generate questions in response to the first question with the aim of using that additional input to answer the question more accurately. You can also specify how many sub-questions the AI tool can ask and provide additional question refinement in the form of context.

Step 3: Instruct the AI tool to combine your answers to the sub-questions it provides to produce the final answer to the main (overall) question. Once the AI tool acknowledges your instructions, ask your question.



Worked Example: [Applying the Cognitive Verifier Pattern](#)

Flipped Interaction Pattern

In this approach, you instruct the AI tool to generate questions for you to answer to enable it to improve the required output.

Example:

I want to create a workshop plan to develop a strategic plan for my organisation for the coming year. You should ask me questions until you have enough information to create the lesson plan. Ask one question at a time.

It consists of two steps:

Step 1: Specify the goal of the interaction, e.g. I want you to ask me questions to achieve X (replace X with the desired goal). Focus on a particular topic or outcome.

Step 2: Specify the duration of the interaction and/or the number of questions it can ask, e.g. You should ask me questions until you have sufficient information to produce the desired outcome.

After inputting the above steps, prompt the AI tool to ask you the first question, e.g. “Ask me the first question.”

This pattern can be a useful tool for getting an AI tool to provide you with your required output based on questions it asks you. While similar to the Cognitive Verifier pattern (in which the AI tool is also asked to generate questions), it places more onus on the AI tool to ask you clarifying questions until it can generate the required output.



Worked Example: [Applying the Flipped Interaction Pattern](#)

Patterns which use a standard template

Template Pattern

This method constitutes an attempt to control the structure and context of the AI tool output by providing the data and specifying a template to be used to format the information provided. You need to know, beforehand, what formatting components your AI tool will recognize to be able to respond appropriately/accurately.

A template pattern exchange contains the following steps:

Step 1: Specify about a template, e.g. “I am going to provide a template for your output.”

Step 2: Indicate the format of the placeholder you will use. This must be something that the AI tool has been trained on, such as using ALL CAPS, enclosing in brackets/markup tags, etc. For example, “CAPITALIZED WORDS are my placeholders.”

Step 3: Instruct the AI tool to fit the output into one or more of the placeholders provided, e.g., “Fill in my placeholders with your output.”/ “Any time that you generate text, try to fit it into one of the placeholders that I list.”

Step 4: Instruct the AI tool to not deviate from the template by trying to include other elements, e.g. “Please preserve the formatting and overall template that I provide.” While this is a useful

way to control the output of the AI tool, a potential drawback of this pattern is that specified template components may exclude potentially useful information.

Best for: Tasks that require following a fixed structure or format for the output, such as generating computer code, specific poetic styles or technical reports.

 Worked Example: [Applying the Template Pattern in ChatGPT](#)

There are many other types of prompts that you can explore, such as prompts that use keywords, provide detailed examples, partial input, instructions, or contextual information.

Take away points

It is worth spending time developing your queries or prompts to get the best results. Try more than one type of prompt and consider using different AI tools depending on what you are seeking so you can get to know their strengths.

Assume you will have multiple interactions until you get the desired result. Keep refining your prompt based on the responses you received previously and keep asking for further refinements. You can include many conditions and specific qualifiers in your questions. Your approach should be iterative so that each set of output improves on the previous one. Use your existing knowledge to tailor each request to make it as specific and relevant as possible (White, 2023c).

It is also possible to combine patterns, and you should experiment to see what gets you the best results. Here is one example:

“From now on, whenever I ask a question, ask four additional questions that would help you produce a better version of my original question (Cognitive verification pattern). Then, use my answers to suggest a better version of my original question (Question refinement pattern). After the follow-up questions, temporarily act as a user with no knowledge of AWS and define any terms that I need to know to accurately answer the questions (Role playing pattern).” (White et al., 2023: 10, our annotation)

Be aware of the in-built bias which can skew your results, particularly that the training data will be largely based on content available on the open internet which is dominated by mainstream

knowledge and languages. To control for the bias, give explicit instructions and use consciously inclusive language.

The following illustrates the power of prompts to skew outputs and demonstrates the importance of refining and evaluating first responses.

First Prompt	Provide a list of the most successful entrepreneurs of all time.
Bard response	Generates a list of 10 business leaders from the USA (Andrew Carnegie, John Rockefeller, Thomas Edison, Oprah Winfrey, Sam Walton, Charles Schwab, Steve Jobs, George Bezos, Bill Gates, Elon Musk)
Follow up prompt	This list comprises people from the USA only. Please revise the list to be more globally representative.
Bard	Provides a list with 10 people mostly from Asian countries (Jack Ma, Masayoshi Son, Bernard Arnault, Amancio Ortega, He Xiangjian, Lei Jun, Ren Zhengfei, Ma Huateng, Lee Byung-chul, Chung Ju-yung, Azim Premji)
Additional follow up prompt	Review the first and second lists you produced and provide a new list of the top 10 entrepreneurs globally, based on ranking the success of the companies they are associated with.
Bard	Here is a list of the top 10 most successful entrepreneurs globally, based on the success of their companies: (Elon Musk; Jeff Bezos; Bill Gates; Mark Zuckerberg; Jack Ma; Masayoshi Son; Bernard Arnault; Amancio Ortega; He Xiangjian; Lei Jun)

While it is possible to provide the AI tool with information that may assist it to generate the required output, remember that this information could also be incorporated into the server on which the tool runs and then may become accessible to other users of that tool (White, 2023a).

Always view Generative AI as a tool that can offer assistance, never an author. Take responsibility and critically review any outputs before using them.

Addendum: AI Prompt Pattern Examples

Role-Playing (Persona) Patterns

Prompt

Act as a 3D modelling teacher who uses Blender 3D. A student is struggling to conceptualize procedural nodes-based 3D modelling. Provide a metaphor for understanding the concept of nodes-based 3D modelling and help the student understand when to use traditional modelling and when to use procedural modelling. Include some points on getting started with some simple nodes as a way of getting started successfully.

Bing chat



Hello, this is Bing. I can help you with learning 3D modelling using Blender 3D.

One way to understand the concept of nodes-based 3D modelling is to think of it as a recipe for creating a 3D object. A recipe consists of ingredients and instructions that tell you how to combine them to make a dish. Similarly, a nodes-based 3D model consists of nodes and links that tell you how to manipulate them to create a shape.

Nodes are like the ingredients of the recipe. They are the basic building blocks that can store or generate data, such as colors, textures, coordinates, vectors, etc. Links are like the instructions of the recipe. They are the connections that transfer data from one node to another, creating a flow of information.

By using nodes and links, you can create complex and realistic 3D models without having to manually sculpt or edit every detail. You can also easily change or tweak any aspect of the model by adjusting the nodes or links, without affecting the rest of the model.

Some advantages of using nodes-based 3D modelling are:

It allows you to create procedural models, which are models that are generated by an algorithm rather than by hand. Procedural models can have infinite variations and details, and can be easily customized or randomized.

It saves you time and memory, as you don't have to store or load large amounts of geometry or texture data. You only need to store or load the nodes and links that define the model.

It gives you more control and flexibility, as you can modify any part of the model by changing the nodes or links, without affecting the rest of the model.

Some disadvantages of using nodes-based 3D modelling are:

It has a steep learning curve, as you need to understand how different types of nodes work and how to connect them properly.

It can be difficult to achieve precise or specific results, as you have to rely on the algorithm rather than your own intuition or vision.

It can be slow or unstable, as it requires more computation power and memory to process the nodes and links.

When to use traditional modelling and when to use procedural modelling?

There is no definitive answer to this question, as it depends on your personal preference, style, and goals. However, some general guidelines are:

Use traditional modelling when you want to create simple or stylized models, such as low-poly or cartoon-like models. Traditional modelling gives you more direct control and feedback over the shape and appearance of your model.

Use procedural modelling when you want to create complex or realistic models, such as landscapes, plants, rocks, clouds, etc. Procedural modelling allows you to generate detailed and varied models with less effort and time.

How to get started with some simple nodes?

A good way to get started with nodes-based 3D modelling is to use the Geometry Nodes Editor in Blender 3D. The Geometry Nodes Editor is a tool that lets you create and edit nodes-based 3D models in a visual way.

To use the Geometry Nodes Editor, you need to follow these steps:

Create a new Geometry Nodes Modifier on an object in your scene. This will open the Geometry Nodes Editor window.

In the Geometry Nodes Editor window, you will see two default nodes: a Geometry Node and a Group Output Node. The Geometry Node represents the input geometry of your object, and the Group Output Node represents the output geometry of your modifier.

To add new nodes, press Shift + A and select a node from the menu. You can choose from different categories of nodes, such as Input, Output, Attribute, Geometry, Mesh, Point, etc.

To connect nodes, drag a link from an output socket (the circles on the right side of a node) to an input socket (the circles on the left side of a node). You can also cut links by pressing Ctrl + Left Click on them.

To edit node properties, select a node and look at the sidebar on the right side of the window. You can change various parameters such as values, colors, vectors, etc.

To preview your model, look at the 3D viewport window. You can also toggle between solid and wireframe mode by pressing Z.

Here is an example of a simple nodes-based 3D model that creates a sphere with random spikes:

Geometry Node -> Point Distribute Node -> Point Instance Node -> Group Output Node

Point Distribute Node: Density = 1000

Point Instance Node: Geometry = Icosphere

Icosphere: Radius = 0.1

I hope this helps you understand the concept of nodes-based 3D modelling better. If you have any questions or feedback, please let me know.

Here are some other role-playing pattern prompt examples for various subjects:

Task	Example of prompt
Social sciences lecturers can use generative AI tools to construct historical role play activities. The same could be done for historical science figures.	Hello ChatGPT, I am Nelson Mandela. I would like to discuss my political role in the anti-apartheid struggle.
Students can be assigned a task that requires them to have a conversation with a Chatbot in the persona of that character.	Can you play the role of Elizabeth Gilbert from Eat, Pray Love? I want to know more about your thoughts on self-discovery and love.
Students can role play as scientists conducting an experiment in a lab with specific roles can be given such as a researcher and lab technician.	Can you play the role of a Chemist who is testing and comparing the density of water and ethanol?
Chat bots can generate mathematical and science scenarios such as solving real-world math problems.	Can you play the role of my mathematics teacher and create a set of prompts that will help me to design a propeller boat using mathematical concepts? Don't give me the final answer but correct me as I move through the steps.
Chat bots can generate real life type of scenarios where students communicate in a foreign language such as asking for directions to a stranger.	Can you please act as my kiSwahili tutor and tell me how I can ask to buy tickets for a theatre show?

Role-Playing (Persona) Pattern (Bard, Bing, and ChatGPT)

Bard	Bing	ChatGPT
Prompt: You are an object-oriented programming lecturer.		

Introduce students to Python programming language.

The same prompt (displayed above) requesting the tool to act as a lecturer was entered into three generative AI tools and the following outcomes were noted:

Bard provides different options for the answer in that three are immediately generated and can be toggled between. The style in which the answer is provided is much like what one might find in a textbook.

Bing's output has an interactive feel to it. It addresses the students in the way in which a lecturer would, provides an example that an explanation is included with and speaks to what will be covered in the next chapter.

The output provided by ChatGPT is a combination of the styles of Bard and Bing in that it addresses the students in the way in which a lecturer would, but the content is written in a textbook style.

Both Bard and Bing caution the user with regard to the generated code.

All three tools followed the prompt without issue, but it would be the user's decision to decide on the preferred output style and refine the interaction accordingly. The outputs can and should be edited with further prompts.

Audience Pattern (ChatGPT 3.5)

Prompt

Explain information retrieval to me summarised in a set of bullet points. Assume I am a university mathematics lecturer.

Chat GPT



ChatGPT

Certainly! Here's a summary of information retrieval in a set of bullet points:

Definition: Information retrieval (IR) is a field in computer science and information science that focuses on the process of obtaining relevant information from large collections of data, typically text-based.

Key Goals:

Retrieve relevant documents or data in response to user queries.

Evaluate and rank results based on relevance.

Support efficient and effective access to information.

Components:

Document Collection: The set of documents, such as web pages, books, or articles, that the system searches through.

Query: The user's request for information, usually expressed as a set of keywords or a longer natural language query.

Indexing: The process of pre-processing and structuring documents for faster retrieval, often involving techniques like tokenization and stemming.

Ranking: Determining the order in which documents are presented to the user based on their relevance to the query.

Methods:

Boolean Retrieval: Uses logical operators (AND, OR, NOT) to match documents based on the presence or absence of keywords.

Vector Space Models: Represent documents and queries as vectors in a multi-dimensional space, allowing for similarity-based ranking.

Probabilistic Models: Rank documents based on the likelihood of relevance given the query.

Machine Learning and Deep Learning: Utilize advanced algorithms, including neural networks, to improve retrieval and ranking.

Evaluation Metrics:

Precision: Measures how many of the retrieved items are relevant.

Recall: Measures how many of the relevant items were retrieved.

F1 Score: A combination of precision and recall for a balanced evaluation.

Mean Average Precision (MAP): A common metric for ranked retrieval systems.

Challenges:

Scalability: Managing and searching through vast amounts of data efficiently.

Relevance Ranking: Developing accurate models for ranking documents.

Query Understanding: Interpreting user queries effectively, including handling synonyms and natural language queries.

Applications:

Web Search Engines: Google, Bing, etc.

Digital Libraries: Managing and searching through vast collections of documents.

Recommendation Systems: Suggesting products, content, or services to users.

Information Filtering: Filtering emails, news articles, or social media content.

Mathematical Foundations: Information retrieval often involves mathematical concepts like set theory, probability theory, and linear algebra for indexing and ranking.

Ongoing Research: Continual advancements in IR include deep learning techniques, personalized search, and handling multimedia data.

Relevance to Mathematics: Information retrieval intersects with mathematics through the use of mathematical models, algorithms, and statistics for improved search and retrieval processes.

Ethical Considerations: Concerns about privacy, bias, and fairness in information retrieval systems are significant topics of discussion and research.

References: Numerous academic journals, conferences, and research groups are dedicated to advancing the field of information retrieval.

As a mathematics lecturer, you may find relevance in the mathematical underpinnings and modeling techniques used in information retrieval, especially if you're interested in data analysis and algorithms.

Here are some other audience pattern prompt examples:

Task	Example of prompt
Explaining concept	a Explain large language models to me. Assume that I know nothing about Computer Science.
Generate outlines for writing assignments	I have to write an essay about the impact of plastic use on the environment for my environmental science course. Please help me create an outline to map out my ideas. Assume I am a first-year undergraduate university student.

Question Refinement Pattern – Exchange 1 (ChatGPT-3.5)

Prompt

From now on, whenever I ask a question, suggest a better question, and ask me if I would like to use it instead.

Chat GPT



ChatGPT

Certainly! I can suggest better questions when I think there's room for improvement. However, I'll always prioritize your original question and provide the suggested question as an option. If you'd like to use the suggested question instead, simply let me know, and I'll proceed with that. Please go ahead and ask your first question or request, and I'll do my best to assist you.

Additional follow up prompt

Should I enroll at the University of Cape Town to study?

ChatGPT

Here's a suggested rephrased question: "What factors should I consider when deciding whether to enroll at the University of Cape Town for my studies?"

Would you like to use this suggested question instead, or do you have any specific criteria or preferences you'd like me to consider when providing advice?

Additional follow up prompt

Yes please use that question

ChatGPT

Of course! Here's your revised question: "What factors should I consider when deciding whether to enroll at the University of Cape Town for my studies?"

ChatGPT

Choosing a university for your studies is a significant decision, and there are several factors you should consider when deciding whether to enroll at the University of Cape Town (UCT) or any other institution. Here are some important factors to take into account:

Academic Programs:

Ensure that UCT offers the academic programs and majors that align with your interests and career goals.

Research the reputation and rankings of specific programs or departments relevant to your field of study.

Accreditation:

Verify that UCT is accredited by the appropriate educational authorities to ensure the validity of your degree.

Location:

Consider the location of the university. UCT is in Cape Town, South Africa, which is known for its natural beauty and cultural diversity. Think about whether you'd enjoy living in this location.

Cost and Financial Aid:

Calculate the cost of tuition, fees, and living expenses, and assess whether you can afford it.

Explore scholarship and financial aid opportunities that UCT may offer.

Reputation and Rankings:

Research the university's reputation and rankings both nationally and internationally.

Read reviews and seek the opinions of current and former students.

Campus Facilities:

Visit the campus or research the available facilities, including libraries, laboratories, sports facilities, and student housing.

Faculty and Research:

Investigate the faculty's credentials and their expertise in your chosen field.

Look into the university's research initiatives and opportunities for student involvement.

Student Life:

Consider the extracurricular activities, clubs, and organizations available on campus.

Think about the social and cultural opportunities offered in and around Cape Town.

International Opportunities:

Assess the availability of exchange programs or study abroad opportunities, as they can enhance your educational experience.

Career Services:

Investigate the university's career services, internships, and job placement support.

Look for connections with local and international companies and organizations.

Diversity and Inclusivity:

Consider the diversity of the student body and the university's commitment to inclusivity and equal opportunity.

Campus Safety:

Research the safety of the campus and the surrounding area.

Alumni Network:

A strong alumni network can be beneficial for networking and job opportunities. Research the reach and influence of UCT's alumni.

Personal Preferences:

Consider your personal preferences, such as class sizes, teaching style, and the overall atmosphere of the university.

Long-Term Goals:

Think about your long-term career goals and how UCT can help you achieve them.

Admission Requirements:

Ensure you meet the university's admission requirements and deadlines.

It's crucial to visit the campus if possible, talk to current students and faculty, and carefully review the university's website and materials to gather the information you need to make an informed decision about enrolling at the University of Cape Town or any other institution. Each student's priorities and needs are different, so weigh these factors based on your specific circumstances and aspirations.

Prompt

Whenever I ask a question about studying at the University of Cape Town, suggest a better question which emphasises the academic merits of the institution. Ask me for the first question to refine and ask me if I would like to use that question instead.

Chat GPT



ChatGPT

Of course, I can help you refine questions about studying at the University of Cape Town to emphasize its academic merits. Please provide me with the first question you'd like to refine, and I'll suggest a better question for you. If you prefer to use the refined question, just let me know.

Follow up prompt

Should I enroll at the University of Cape Town to study?

ChatGPT

To emphasize the academic merits of the University of Cape Town, how about this refined question: "What academic programs and research opportunities make the University of Cape Town a compelling choice for prospective students?" Would you like to use this question instead?

Follow up prompt

Yes, answer the question

ChatGPT

Certainly! The University of Cape Town (UCT) is known for its strong academic programs and research opportunities. UCT offers a wide range of undergraduate, postgraduate, and doctoral programs in various fields, including humanities, sciences, engineering, business, health sciences, and more. Some key academic merits of UCT include:

World-Class Faculty: UCT has a renowned faculty consisting of experts, scholars, and researchers who are leaders in their respective fields.

Research Excellence: UCT is a research-intensive institution, and its research output is highly regarded globally. It has a significant impact on the advancement of knowledge in numerous disciplines.

State-of-the-Art Facilities: The university provides modern, well-equipped facilities, laboratories, and libraries to support students and researchers.

International Reputation: UCT consistently ranks as one of the top universities in Africa and is well-regarded worldwide.

Diverse Student Body: UCT attracts a diverse student body, fostering a rich cultural and academic exchange environment.

Global Opportunities: The university offers international exchange programs, encouraging students to broaden their horizons by studying abroad.

Community Engagement: UCT is actively involved in community engagement and outreach, providing opportunities for students to make a positive impact.

If you have specific questions about a particular academic program or department at UCT, feel free to ask for more detailed information.

Cognitive Verifier Pattern (ChatGPT-3.5)

Prompt

When I ask a question, follow these rules. Generate a number of additional questions that will allow you to more accurately answer the question. After three additional questions, combine the individual answers to respond to the original question.

Chat GPT



ChatGPT

Sure, I can follow those rules to help provide more accurate answers to your questions. Please go ahead and ask your original question, and I'll generate up to three additional questions to gather more information before providing a response.

Follow up prompt

How many South Africans live without indoor sanitation?

ChatGPT

To answer your question accurately, I'll generate up to three additional questions to gather more information:

Do you want to know the current number of South Africans living without indoor sanitation, or are you interested in a specific year or time frame?

Are you looking for statistics related to urban or rural areas, or an overall national figure?

Could you specify what you mean by "indoor sanitation"? Does this include access to flush toilets, improved pit latrines, or other forms of indoor sanitation facilities? Once I have this additional information, I can provide a more precise response.

Follow up prompt

1. Could you provide a percentage of households currently living without indoor sanitation? 2. overall national figure 3. indoor sanitation = piped water and flush toilets

ChatGPT

Thank you for providing the additional questions. Here are the answers based on the information available up to my last knowledge update in January 2022:

Percentage of households currently living without indoor sanitation in South Africa: As of my last knowledge update, it was reported that approximately 19% of South African households did not have access to basic sanitation services, which includes both piped water and flush toilets. However, please note that these statistics may have changed since then, so it's essential to refer to more recent data for the most accurate figures.

Overall national figure: The overall national figure for the percentage of South African households without indoor sanitation was around 19% according to the information available up to 2022. To get the most up-to-date statistics, you should consult the latest reports or government sources.

Indoor sanitation - piped water and flush toilets: In the context of access to indoor sanitation, "indoor sanitation" typically includes access to piped water and flush toilets. As of my last knowledge update, efforts were being made to improve access to these facilities in South Africa. Please refer to more recent sources for the latest developments and statistics related to indoor sanitation in South Africa.

Keep in mind that statistics can change over time, so I recommend consulting recent reports or government sources for the most current information on this topic.

Flipped Interaction Pattern (ChatGPT-3.5)

Prompt

I want to create a lesson plan on climate change for a South African Grade 8 class. You should ask me questions until you have enough information to create the lesson plan. Ask one question at a time.



ChatGPT

Certainly! Let's start by gathering some information for your Grade 8 climate change lesson plan in South Africa. What is the typical duration of a class session for Grade 8 in South Africa?

Follow up prompt

Usually 45 minutes but sometimes there are double sessions

ChatGPT

Thank you for the information. Now, let's consider the following:

What specific learning objectives or goals would you like to achieve with this lesson on climate change for Grade 8 students in South Africa?

Follow up prompt

For the class to understand the basic causes of human induced climate change and likely impacts in this region

ChatGPT

Great! Understanding the causes of human-induced climate change and its likely impacts in the South African region will be the focus of your lesson.

Do you have any specific resources or materials available for this lesson, such as textbooks, videos, or online resources related to climate change that you would like to use?

Follow up prompt

Online resources from the Keep it Cool Climate Change Education Project and the IPCC

ChatGPT

Excellent! You have access to online resources from the "Keep it Cool Climate Change Education Project" and the Intergovernmental Panel on Climate Change (IPCC).

Follow up prompt

How many class sessions or lessons do you have available to cover this topic, considering that some sessions may be double sessions?

ChatGPT

Thank you for providing the number of class sessions. Now, let's plan the structure of your lesson.

Would you like to start with an introductory session and then proceed with more in-depth lessons, or do you have a different structure in mind for these 5 sessions?

Follow up prompt

I would like to start with an excursion somewhere within 5kms of Cape Town city that could be used to highlight the issues as an impactful start to the learning process - then proceed to the science and political issues.

ChatGPT

Starting with an excursion near Cape Town to highlight the issues related to climate change sounds like an engaging way to kick off the lesson. Do you have a particular location or site in mind for this excursion, or would you like some suggestions for suitable places within 5km of Cape Town city?

Follow up prompt

Yes, let's plan a research project on the impact of extreme weather events at the Sea Point Promenade (damaged by recent storms). What do you suggest for the remainder sessions that can link to that topic and lead into the other areas I need to cover.

ChatGPT

Great choice! Investigating the impact of extreme weather events at the Sea Point Promenade is a meaningful project. Here's a suggested structure for your lesson plan with subsequent sessions:

Session 1: Excursion (3 hours)

Visit Sea Point Promenade, document storm damage, and gather initial data.

Discuss the students' observations and findings during the excursion.

Session 2: Introduction to Climate Change and Extreme Weather (45 minutes)

Provide an overview of climate change and its link to extreme weather events.

Discuss the concept of attribution - how scientists determine if an extreme weather event is linked to climate change.

Show relevant resources from the "Keep it Cool Climate Change Education Project."

Session 3: The Science of Climate Change and Global Warming (45 minutes)

Focus on the science behind climate change, including the greenhouse effect, the role of greenhouse gases, and global warming.

Explain how human activities, such as burning fossil fuels, contribute to increased greenhouse gas concentrations.

Explore the basic physics of the Earth's energy balance.

Use visuals and simple experiments to illustrate these concepts.

Session 4: Analyzing Data (45 minutes)

Have students work on their research project: Analyze data from the Sea Point Promenade excursion and discuss the scientific aspects related to the storm damage and climate change.

Session 5: Regional Impacts, Adaptation, and Climate Policy (45 minutes)

Explore the regional impacts of climate change on Cape Town and South Africa.

Discuss how communities are adapting to extreme weather events and the role of local government.

Introduce international agreements and policies related to climate change.

Encourage a class discussion on what can be done to combat climate change.

Incorporating a dedicated science session (Session 3) will help students gain a deeper understanding of the scientific principles underpinning climate change and global warming. It's essential to link the scientific aspects to the practical observations made during the excursion and the broader context of regional impacts and climate policy.

Template Pattern (ChatGPT-3.5)

Prompt

I am going to provide a template for your output. CAPITALIZED WORDS are my placeholders.

Fill in my placeholders with your output. Please preserve the overall formatting of the template.

My template is:

****Question:***QUESTION*

Answer:ANSWER

Use the data I will give to you below and create 5 questions and answers using the template I provided above.

Data = """*Prompt engineering, primarily used in communication with a text-to-text model and text-to-image model, is the process of structuring text that can be interpreted and understood by a generative AI model.[1][2] Prompt engineering is enabled by in-context learning, defined as a model's ability to temporarily learn from prompts. The ability for in-context learning is an emergent ability[3] of large language models.*"""

A prompt is natural language text describing the task that an AI should perform.[4] A prompt for a text-to-text model can be a query such as "what is Fermat's little theorem?"[5] a command such as "write a poem about leaves falling",[6] a short statement of feedback (for example, "too verbose", "too formal", "rephrase again", "omit this word") or a longer statement including context, instructions,[7] and input data. Prompt engineering may involve phrasing a query, specifying a style,[6] providing relevant context[8] or assigning a role to the AI such as "Act as a native French speaker".[9] Prompt engineering may consist of a single prompt that includes a few examples for a model to learn from, such as "maison -> house, chat -> cat, chien ->",[10] an approach called few-shot learning.[11]

When communicating with a text-to-image or a text-to-audio model, a typical prompt is a description of a desired output such as "a high-quality photo of an astronaut riding a horse"[12] or "Lo-fi slow BPM electro chill with organic samples".[13] Prompting a text-to-image model may involve adding, removing, emphasizing and re-ordering words to achieve a desired subject, style,[1] layout, lighting,[14] and aesthetic.



ChatGPT

Question: What is the primary purpose of prompt engineering in communication with generative AI models? Answer: The primary purpose of prompt engineering in communication with generative AI models is to structure text that can be understood by the AI model, enabling it to perform specific tasks.

Question: How is in-context learning defined in the context of prompt engineering? Answer: In the context of prompt engineering, in-context learning is defined as a model's ability to temporarily learn from prompts, allowing it to adapt and generate appropriate responses.

Question: Can you provide an example of a prompt for a text-to-text model? Answer: Certainly, a prompt for a text-to-text model could be a query like "what is Fermat's little theorem?" or a command like "write a poem about leaves falling."

Question: What are some elements that prompt engineering may involve when communicating with a text-to-text model? Answer: Prompt engineering when communicating with a text-to-text model may involve phrasing a query, specifying a style, providing relevant context, or even assigning a specific role to the AI, such as "Act as a native French speaker."

Question: How does prompting work when communicating with a text-to-image or text-to-audio model? Answer: When communicating with a text-to-image or text-to-audio model, prompting involves providing a description of the desired output, such as requesting "a high-quality photo of an astronaut riding a horse" or specifying the desired audio style like "Lo-fi slow BPM electro chill with organic samples."

Here is [simple prompt template](#) you could start with which combines some of the other prompt patterns:

Act as {{persona you want the AI tool to simulate}}.

I am {{audience}}.

I want you to {{task}}.

Constraints: • {{constraint 1}} • {{constraint 2}} • {{constraint 3}}

Output format: • {{format instruction 1}} • {{format instruction 2}}

1.5. Student Guide: Ethical use of generative AI for research purposes

The availability and rapid uptake of Artificial Intelligence (AI) tools poses both challenges and opportunities for researchers. This guide looks at how generative AI is starting to be used, the implications for research integrity, and suggestions on how these tools can be used productively and ethically for the research process.

The availability and rapid uptake of Artificial Intelligence (AI) tools poses both challenges and opportunities for researchers. This guide looks at how generative AI is starting to be used, the implications for research integrity, and suggestions on how these tools can be used productively and ethically for the research process. Examples of interest to researchers, students, supervisors, and others are included.

While the outputs produced by generative AI are widely deemed unacceptable as substitutes for what researchers are expected to perform, there are routine tasks forming a part of research processes where generative AI is being used. Researchers, when faced with identifying relevant articles, have for example used generative AI to summarise these and make the task of reading selected articles more manageable. Ethical concerns may arise if, for example, sensitive interview transcripts needed to be summarized, as uploading these into generative AI may potentially give others access to them and violate confidentiality.

Balancing productivity and ethical concerns requires careful consideration. This is a changing and uncertain space requiring those involved to develop shared understandings. This guide provides some cases and examples that can support this process.

General use

This guide deals primarily with generative AI such as ChatGPT, Gemini, Copilot Chat, and Claude, which are based on large language models. This natural language styles of generative AI invites one to have a conversation, however these are not human conversations nor is the

information necessary factually correct. The large language models that are responding to your prompts by predicting a likely next word. By crafting prompts one can control and direct what is generated so that it will be more useful.

Providing context to generative AI will often improve the response. For example, it may be useful to include your role, your knowledge area, or any constraints on the output. In ChatGPT for example, you can include this under “Custom Instructions” that are associated with your login (bottom left corner on the ChatGPT page). These would be used each time you use ChatGPT. Otherwise include such information when you start your session to help make the results more relevant.

Additionally, you can request generative AI to follow the preferred style of writing, by providing some examples.

Prompt: *"Please help me write helpful information for university researchers. I want you to understand this writing style so I will provide some examples. Remember my writing style as being MY_STYLE. Afterwards when I ask you to write using MY_STYLE, then follow this style of writing."*

Examples other than academic writing might include computer programming, and marketing applications where the expectations would be different. More generic strategies for using Generative AI are introduced in other CILT Guides (see A Generative AI Primer). Specialized generative AI tools have been developed for some of these tasks. Examples of these are listed later in this guide as part of a research cycle.

General characteristics

Knowing how generative AI functions can help with crafting prompts and anticipating cases where it will be useful:

Language generation: Generative AI models are trained on natural language text and can generate coherent and grammatically correct content. Useful for tasks like summarising articles, drafting papers, or generating ideas.

Large-scale information base: Generative AI models are trained on many diverse sources, which provides them with a broad information base. Researchers could use these models to gain insights into various subjects, identify relevant literature, and explore interdisciplinary connections.

Intended use cases: Typically, generative AI models are trained to conduct conversations in a customer support role, so are polite, direct and avoid sensitive topics. There are many commercial applications for this chatbot use case. The support role can be established, to play the role of editor, tutor, or idea generator. There are opportunities for follow-up questions and requesting revisions to improve responses.

Speed and efficiency: Generative AI models can quickly process large amounts of new data and generate outputs immediately. Researchers could save time and effort on tasks such as identifying possible sources for literature reviews, in getting assistance with data analysis, or brainstorming to get started on writing tasks.

Pattern recognition: Generative AI models are capable of identifying patterns, trends, and relationships within data. Generating code to perform additional analysis and outputs on your data is also possible, such as in Python or R. Valuable for discovering new insights or generating hypotheses in research.

Customizability: Researchers can tailor generative AI models to their specific needs. Generative AI can be asked to adopt specific genres for creating a variety of outputs suitable to different audiences. Researchers can fine-tune models.

Considerations and limitations

Being aware of the considerations and limitations helps identify where unreliable responses are more likely and where checking is needed.

Lack of transparency: The inner workings of many generative AI models are complex and opaque, making it difficult to understand how they generate their outputs. Can be challenging for researchers to be transparent about their process. The lack of transparency makes it difficult to identify errors in the generated data.

Inherent biases or errors in training dataset: Generative AI models learn from the data they are trained on, which may contain bias or particular perspectives. Researchers need to critically evaluate outputs for bias/hegemonic perspectives.

Misuse of generated data: Generative AI can also be used to create fake data or manipulate existing data. Researchers need to critically evaluate any data.

Reproducibility issues: Generative AI models may not be reproducible, as the same model may produce different outputs each time it is run due to the random nature of the generative process. This can make it difficult for researchers to replicate results and make claims about reproducibility.

Research integrity implications

Researchers and students should follow the widely discussed integrity and ethical issues:

Journals and publications: AI will not be acceptable for assisting with producing most parts of research publications. Some journals will allow AI for conceptualisation or copyediting but not for discussion or conclusions (original contribution). AI can also not be listed as a co-author by most journals. Most journals have clear instructions on how to acknowledge the use of generative AI in research, which differ from disciplines to discipline. Where AI is legitimately used to generate text, this should be referenced (e.g., APA style) with further declarations (i.e., including the prompts used). See for example citing [ChatGPT](#) and [generative AI](#).

Publisher	Authorship policy update	Remark
Springer-Nature (2023)	LLMs, such as ChatGPT, do not satisfy the authorship criteria. However, if the researchers use these tools, s/he must mention their use in the appropriate section of their academic paper, such as the ‘methodology’ or ‘acknowledgements’ section.	Generative AI cannot be an author or co-author. If used this should be clarified in the appropriate section.
Taylor & Francis (2023)	Authors must be accountable for their research work per the publishing agreement. As AI tools do not take this accountability, thus: AI tools cannot be co-authors in an academic paper. However, if a researcher uses these tools, s/he must mention their use in the appropriate section.	Generative AI cannot be an author or co-author. If used this should be clarified in the appropriate section.
Elsevier (2023)	Though AI and AI-assisted technologies help you to enhance the quality and readability of the language of the	Generative AI cannot be an author or co-author.

	work, they do not replace key researchers. Thus, the researchers are not allowed to list AI and AI-assisted technologies as an author or co-author nor cite AI as an author.	AI cannot be cited as an author.
--	--	----------------------------------

Authorship policies of three large publishers (from [Rahman et al 2023](#), p.3)

Students: At the start of any research investigation or course, make clear to students whether generative AI tools can be used and how it can be used. Explain the consequences of academic dishonesty and inappropriate use of AI. Where appropriate, include conditions in the plagiarism declaration/honour pledge or in the assignment submission instructions. Have a discussion with your students about what ethical AI use would look like. Lance Eaton's '[Syllabi Policies for Generative AI](#)' is a collection of statements used in various universities that provide some examples. There is further guidance to students on using AI in another CILT [student guide](#).

Data privacy: Be aware that data inputted into for example ChatGPT3.5 is currently being saved, which could lead to privacy issues. ChatGPT4.0[\[DG1\] \[SW2\]](#) allows you to change this setting (uncheck the setting that data is saved).

Data analysis and consent: You might have to change consent forms to ask participants for permission to use their data for analysis in generative AI.

More broadly, the UCT [Senate Ethics in Research Committee \(EiRC\) Guidelines and recommendations for the use of generative artificial intelligence \(AI\) tools in research](#) establishes values for guiding research integrity:

Honesty in all aspects of research

Professional courtesy and fairness in working with others

Good stewardship of research on behalf of others

Transparency in conducting research and dissemination of findings

Fair practice from conception to implementation of research

Shared accountability in the conduct of research

Indigenous knowledge recognition and epistemic justice

Research process stages with generative AI examples

For each phase in a typical research process may be using specialized generative AI tools. The research phases and examples of these are explored below.

Research process stage	Generative AI examples
Conceptualisation/ ideation	ChatGPT for brainstorming, writing conference abstracts, writing proposals, brainstorming/rephrasing research questions,
Literature review	LitMap or inciteful.xyz to create literature maps around concepts, Typset.io to summarize articles https://scite.ai/ (paid for tool) for analysing references
Data collection	Otter.io to transcribe interviews, summarize interviews, create keywords
Data analysis	ChatGPT or voyant for thematic analysis (ChatGPT 4.0 allows an advanced data analysis)
Write up	Quillbot to rewrite paragraphs, ChatGPT to convert bullet points into paragraphs, to polish writing
Dissemination	Use ChatGPT to rewrite your academic papers, i.e., to summarise or make academic papers more accessible.

Principles for the case studies

Based on experiences at UCT

Tools that are generally used in our research community

Based on both lit review and real-life examples

Caveat: all these tools change all the time and information can be outdated

Stage of research: Conceptualisation and research idea generation

Tool used/links	Descriptions of use	Opportunities / Concerns
-----------------	---------------------	--------------------------

ChatGPT, Gemini, Copilot, Claude and other specialised tools	<p>Brainstorming research ideas and hypotheses: Describe your research area or problem, and associated or potential research questions, approaches, and angles will be produced.</p> <p>Example prompts: How has Soja's Third Space been used in the Student as Partners context?</p> <p>What has not yet been researched in the Student as Partners context?</p> <p>Please help me formulate research questions around student as partners, third space, South Africa, rules of engagement</p>	<p>Quick high-level overview of concepts.</p> <p>Easy to digest / accessible information (often in bullet form)</p> <p>Is like asking a research assistant rather than asking an expert.</p> <p>Need to fact check.</p> <p>Last knowledge update in September 2021</p>
ChatGPT, Gemini, Copilot, Claude	<p>Abstract writing: Create a. initial draft of an abstract or compare one to yours. Provide bullet points and ask it to write as a paragraph.</p>	<p>Helps with starting or rethinking an abstract.</p> <p>Cannot represent your ideas fully.</p>
ChatGPT, Gemini, Copilot, Claude	<p>Problem statement and rationale: Once you have a research topic, ask for help to formulate a problem statement or rationale.</p>	<p>Follows typical format.</p> <p>What is generated is likely hypothetical and without any references. So provides a generalized gap without having accessed any original articles.</p>

Copilot, ChatGPT4.0	Accessing the web: Generative AI tools with connections to the web can cite sources, and perform real time search	Output is often less controlled and may appear uncurated.
--------------------------------	--	---

Stage of research: Literature review

Tool used	Descriptions of use	Opportunities / Concerns
ChatGPT, Gemini, Copilot, Claude	Repetitive and tedious structured tasks: Given the appropriate text, can be asked to format references or produce summaries in a specified format.	Is especially useful for specific use cases. Can have errors, so needs critical editing.
Elicit.org	Automate time-consuming tasks: Can assist with searching and summarizing papers, extracting data, and synthesizing your findings	Full paper search, but only searches the semantic web Additional functionality is requires payment (limited to 5000 free credits)

ChatGPT, Gemini, Copilot, Claude and other specialised tools	<p>Literature reviews. Provide specific research topics or questions to generate related search terms for searching literature databases and web searches.</p> <p>Prompts:</p> <p>Can you suggest some keywords and search strategies that I can use to find relevant sources on "Student as Partners" and equity in higher education?</p> <p>Who are the seminal authors in the Student as Partners and equity literature?</p>	<p>Great overview of current trends and debates.</p> <p>Can't provide references, may hallucinate when asked for locally specific information</p> <p>Information is not necessarily correct: If asked for example for scholars on a particular topic in the Global South, scholars are often made up because they may be underrepresented in the training data.</p>
LitMaps (litmaps.com)	<p>Literature maps: Generates a map of relevant articles related to a specific seed article. Show the top citations and references related to a seed article. Selecting any article one can see who the article cites, and which articles are citing it.</p>	<p>Visual view of research field using citations.</p> <p>No made-up references.</p> <p>Some papers are behind paywalls.</p>
scite.ai	<p>Evaluate articles: Helps with evaluating scientific articles by looking at citations. Shows how a publication has been cited by providing the context of the citation and a classification describing whether it provides supporting or</p>	<p>Contextualises article citations.</p> <p>Mostly medical sciences.</p>

	contrasting evidence for the cited claim.	
Typeset.io (also called scispace)	<p>Summarises articles: Allows one to upload PDFs and summarise articles. Additionally it can respond to questions related to articles.</p> <p>Question prompts:</p> <p>What are novel methodologies in these papers?</p> <p>What are unexpected results in these papers?</p> <p>Are there any future research directions or unanswered questions suggested by these papers?</p>	<p>Aims to help understand research papers better.</p> <p>Has some limitations and requires assessing the responses critically.</p>
Consensus.app	Identify findings in literature: A search engine that uses AI to find insights in research articles.	Extracts findings directly from articles.
Semantic reader	Identify key parts of article: Used to understand document structure and merge it with information via tooltips and other overlays. Category labels include Goal, Method, and Result.	<p>Can customise what is shown.</p> <p>Works better for structured articles.</p>

Stage of research: Data collection/transcription

Tool used	Descriptions of use	Opportunities / Concerns
ChatGPT	Interview and survey questions: With some context, generative AI can help develop, evaluate and test interview protocols or survey questionnaires.	Useful for standard types of questions. May not be appropriate for your target group.
Otter.ai	Transcriptions: Allows live edits of transcriptions while listening to the transcript. Supports analyses of texts by for example providing a list of key words/terms used.	Can learn different accents and improves transcription, identifies speakers. Currently only English transcription possible – no other South African languages. Most features require payment. Free version is limited.

Stage of research: Data analysis

Tool used	Descriptions of use	Opportunities / Concerns
-----------	---------------------	--------------------------

ChatGPT4.0 Plus/ Turbo (code interpreter option)	<p>Python code generation: Using the code interpreter option, you can provide data in a file, then ask for Python code to be generated that will analyze the data. The Python code can be executed. Follow-up prompts can be used to refine the analysis.</p>	<p>The saved Python code can be used to analyse other datasets without uploading any data to ChatGPT.</p> <p>ChatGPT Plus/Turbo is subscription based. Claude is a free alternative.</p> <p>Python code should be checked before reporting.</p> <p>Uploading sensitive data would not be appropriate.</p>
ChatGPT4.0	<p>Advanced data analytics: Can analyse large multimodal models. It generates text outputs (natural language, code, etc.) given inputs consisting of interspersed text and images.</p>	<p>While less capable than an expert in many real-world scenarios, it performs well for many professional and academic tasks.</p> <p>May generate harmful advice, buggy code, or inaccurate information.</p>

<u>Vovant</u>	<p>Text analysis: An open-source, web-based application for performing text analysis. It supports scholarly reading and interpretation of texts or corpus.</p>	<p>It can be used to analyze online texts or ones uploaded by users.</p> <p>Prolonged text-loading time and the challenge of gathering information using some visualization tools.</p>
<u>Maxqda</u> and related data analysis tools	<p>Text thematic analysis: Summarise texts, coding of themes, and identifying common themes across your codings.</p> <p>Can help to explain the coded content.</p>	<p>Can choose the language and the length of the summary.</p> <p>Can modify the summaries as needed.</p>
https://cursor.sh/ (based on VScode)	<p>The AI-first Code Editor Build</p> <p>software faster in an editor designed for pair-programming with AI</p> <p>Support with writing code – integrated with GPT4.0</p>	Scaffolds code, reducing time to write code.

Stage of research: Write up

Tool used	Descriptions of use	Opportunities / Concerns
-----------	---------------------	--------------------------

quillbot.com	Writing assistance: Assists with academic research and writing tasks.. Can improve writing style or paraphrase to shorten, provide vocabulary suggestions, and offer alternative word choices.	Helps brainstorming during the writing process. Offers explanations for its suggestions. Can be misused to obscure plagiarism.
wordvice.ai	Writing style: Helps fix spelling, punctuation, and style errors and improves the clarity and flow of your text.	Has explanations as well as tips to improve writing style. Works on all kinds of academic texts.
Grammarly	Grammar and style: Support academic writing by offering spell and grammar checks.	Works in Word and Google Docs, plugin. Free version works well. Can be a bit annoying when it's installed as a plugin.
Jenni.ai	Create content: Given prompts will suggest additional content for articles, websites, etc. Different citation styles can be selected.	Beyond repetitive tasks, may have ethical concerns, depending on how it is used.

Stage of research: Dissemination

Tool used	Descriptions of use	Opportunities / Concerns
-----------	---------------------	--------------------------

jasper.ai	<p>Copywriting: Used for copywriting and marketing campaigns.</p>	<p>Use in research reporting and public engagement.</p> <p>Is a specialised tool to standardise marketing content.</p>
------------------	--	--

Bespoke research tools

Universities are starting to develop their own tools based on generative AI platforms to support research processes. An example is the suite of tools developed by the [Academic Insight Lab](#) based on the YouAI platform. This includes both free and subscription-based tools.

Purpose statement feedback: A research purpose statement is a concise, clear description of the specific goal or aim that a research project seeks to achieve. A well-defined purpose statement serves as a roadmap for the research, ensuring it remains focused and relevant. This AI tool will provide feedback on your purpose statement to ensure it succinctly describes the study population, approach, and setting of the study. <https://youai.ai/ais/purpose-statement-feedback-tool-b50417ed>

Research problem statement feedback: Crafting a compelling research problem statement is often a daunting task, fraught with the complexities of contextualization, identifying knowledge gaps, and establishing significance. Receive personalized feedback powered by AI to ensure your research problem appropriately frames the research problem you plan to address. <https://youai.ai/ais/research-problem-statement-feedback-tool-8d9f5704>

Lit search: designed to help researchers and academics in their literature search process. My job is to assist you in identifying keywords, synonyms, search strings, and more, based on the key concepts and variables of your problem statement <https://youai.ai/ais/litsearch-guide-a73c20ee/use>

Writing diagnostic tool: This AI-powered writing diagnostic tool evaluates academic writing samples, offering a holistic assessment through the lens of SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis. This sophisticated tool adheres to established best practices for academic writing and identifies key areas for improvement to enhance scholars' effectiveness and success in their research communication. <https://youai.ai/ais/writing-diagnostic-2834c729>

Conclusion

Generative AI is being used to provide creative support to researchers, helping with new idea generation, outlining research topics, writing abstracts, etc. Litmaps, for example, and tools specifically designed to support literature reviews provide more consistent and reliable information. Generative AI is also widely used for performing repetitive and tedious tasks, such as summarising large text and identifying key findings from the literature. In such cases one might tolerate some inaccuracy when using generic Generative AI tools.

However, there are greater concerns about generative AI used for writing an academic article and drawing conclusions and where transparency is required (Hosseini et al., 2023). Generative AI tends to produce hypothetical problem statements and research gaps with fake references. Generative AI for data analysis is currently limited and requires checking. Researchers need to show they have control, with generative AI use being limited to a research assistant role (Rhaman et al, 2023).

2. Policy and Guidelines for use of Generative AI

Cairo University

Faculty of Computers and Artificial Intelligence

FCAI Policy and Guidelines for use of Generative AI in
Postgraduate Studies and Research

September 4, 2023

<<Version 1.0>>

Introduction

Advancing knowledge and technology is one of the key objectives of pursuing postgraduate study and conducting innovative research. The mismatch between the speed by which knowledge and technologies evolves compared to that of regulations and policies has raised several concerns in the academic and research communities. Accordingly, it is important to strike a balance between the benefits and risks of adopting and adapting emerging knowledge and technologies.

Generative artificial intelligence (AI) is one of the recent advances that has been evolving much faster than most organizations can comprehend and control.

Generative artificial intelligence (AI) uses machine learning (ML) to respond to prompts inputted by the user. Several technologies have emerged over the last few years including, for instance, the text generator ChatGPT and the image generator DALL-E 2, and Stable Diffusion.

Recently, the use of generative AI in education, and more specifically, in research, has raised a great concern in academia from both ethical and scientific perspectives. Some recent efforts have emerged to attempt to develop policies, regulations, and guidelines for the various academic stakeholders (researchers, students, faculty staff) to ensure that the use of generative AI in academic does not violate the norms and standards that have been long established and enforced.

This document represents the first attempt for the Faculty of Computers and Artificial

Intelligence- Cairo University (FCAI-CU) to develop its policy for using generative AI technologies in teaching and research. In doing so, this living document serves as initial guidelines that will be adopted for the Academic Year 2023-2024. New versions will be issued as the technology matures and we gain more understanding of the benefits and risks of using generative AI in education. This document is based on various guidelines and policies developed at Stanford, Toronto, and Deakin University.

Policy Development Approach

FCAI-CU, as in all international academic institutes, is keen to progressively develop its policy and regulations for using generative AI in teaching and research. As it progresses, the policy should not prevent innovative use and exploration of the potential of generative AI in teaching and research; while at the same time ensures highest academic standards and adherence to FCAI-CU code of ethics and research regulations.

To this end, this document serves as a set of guidelines for using generative AI in teaching and research at FCAI-CU. While this initial version is being adopted at FCAI-CU, regular feedback will be collected from all stakeholders on the use of the guidelines. In addition, we will continue to monitor and analyze polices and regulations that evolve in worldwide academic institutes. Eventually, a new version of this document will be released in the near future followed by continuous improvement cycles to evolve the document as novel technologies evolve over time.

Key Guidelines for Using Generative AI in Teaching and Research

The following are the key initial guidelines for using generative AI in teaching and research in FCAI-CU for postgraduate studies.

Scope of the Guidelines

These guidelines applies to all courses taught at postgraduate level including

diplomas, Master, and Doctorate programs. It applies to coursework, projects, research proposals, and thesis and dissertations, and related scientific publications.

For scientific publications, researchers will need to consult the publisher policy for any further regulations imposed on the use of generative AI and related technologies.

Coursework

- Instructors are free to decide whether to allow or prevent the use of generative AI in their courses. Instructors shall clearly indicate in course syllabus the rules they see appropriate for the use of generative AI in the various activities of the coursework.
- In the absence of clear instructions on the use of generative AI in the coursework, the use of such tools should be treated as an assistance from another person and evaluated based on the regular policies and regulations for such case.
- In all cases, students are responsible for clearly understand the course policy on using generative AI. When they are in doubt, students must ask their instructors for clarification.
- Students should acknowledge the use of generative AI using the suggested citation in Section 3.4 below.

Research Proposals, Thesis, and Dissertations

The bottom-line, any use of generative AI tools at any step in the process of graduate thesis research, writing, and publication must always take place with full transparency and in adherence to the norms and highest standards of research ethics, intellectual property, and accountability. The following key guidelines represent the minimum expectations for graduate researchers at FCAI-CU when using generative AI tools in their research.

- Supervisors must agree in advance how any generative AI tools will be used at any stage of the thesis/dissertations development. Such agreement must be in writing using unambiguous statements and well-defined scope of use. The

original approval document shall be kept in the student file at postgraduate office with clear timestamp log for the document generation.

- If the use of generative AI tools is approved by supervisor(s), it must be clear to the student what evidence they need to provide to demonstrate their own contributions and how they made use of any AI tools, and how their work will be assessed by the supervisor and committee.
- Unauthorized use of generative AI tools for scholarly work shall be treated as research misconduct, and hence, police and regulations for academic research misconduct will be applied.
- Students are fully responsible for contents generated by AI that they decide to include in their thesis. Students must be able to explain and defend any use of generative AI, as well as the contents of the thesis during their thesis/dissertation defense.
- Any use of generative AI tools must be appropriately acknowledged (See Section 3.4 for more details). This includes the use of generative AI tools in searching, designing, outlining, drafting, writing, or editing the thesis/dissertation, or in producing audio or visual content for the thesis/dissertation, and may include other uses of generative AI.
- students working with sensitive types of data (e.g., confidential information related to specific areas or industry partnership, human/individual data) are responsible for the risk of using such data with third party generative AI tools If such data is revealed or disclosed. Students must take appropriate measures to protect the data and/or obtain written approvals from relevant stakeholders before submitting the data to the third party tools.
- Supervisors are highly discouraged to give permissions to their students to use generative AI tools in writing their thesis/dissertation. This is because learning the practices of scientific scholarly writing is crucial for graduate students. The use of generative AI tools could adversely impact the development of these writing skills.

Referencing Guidelines

The following guidelines are adopted from Deakin University guide on referencing:

- Provide details of the owner/publisher of the AI tool and the year of publication.

You might also provide further details of how you used the AI tool, for example a transcript on inputs/outputs, in an appendix.

- References: Author/Owner of AI model. (Year). Name of AI model (Version)

[Type or description of AI model]. URL

- Example OpenAI. (2023). ChatGPT (May 24 version) [Large language model].

<https://chat.openai.com/chat>

3. ChatGPT & other AI tools for Learning and Teaching

This resource is licenced under a CC Attribution Non-Commercial ShareAlike licence

ChatGPT & other AI tools for Learning and

Teaching

Version 1.02 of 26 August 2023

A comprehensive ‘living’ resource about generative and other ‘artificial intelligence’ that will be constantly updated with new resources. To contribute to this resource, please access the google doc version of this resource.

The initial outline for this resource was generated by ChatGPT. The prompt used was:
Create a comprehensive teaching resource about using ChatGPT in higher education.
Include sections on what ChatGPT is, how ChatGPT works, benefits of ChatGPT for Higher Education, how ChatGPT can be used in the classroom, challenges of ChatGPT for higher Education, how students can misuse ChatGPT, and how such misuse by students can be detected. Cite some resources for further reading under every section.

Introduction to ChatGPT

ChatGPT, or "Chat Generative Pre-training Transformer," is a cutting-edge natural language processing (NLP) tool that can generate human-like text based on a given prompt or context.

Developed by OpenAI, ChatGPT is a variant of the GPT-3 model, which has been trained on a massive amount of text data and has the ability to generate text that is often indistinguishable from text written by a human.

What is ChatGPT?

ChatGPT is a language model that uses deep learning techniques to generate text. It is pre-trained on a massive amount of text data, including books, articles, and websites, which allows it to generate text that is similar to text written by humans. It is designed to generate text in response to a given prompt or context, and it can be fine-tuned for specific tasks such as language translation, summarization, and question answering.

Relevant resources

- Something wicked this way comes: ChatGPT, Artificial Intelligence's quantum leap Steven Boykey Sidley, a professor of practice at JBS, University of Johannesburg provides a brief history of the development of AI and describes how CHATGPT is evolving.
- What is ChatGPT? And questions you have been afraid to ask A succinct background to the tools development
- What is Chat GPT Explained Helpful video explaining what Chat GPT is, what it can do and its history as a product of Open AI.

How ChatGPT works

ChatGPT uses a technique called "transformer architecture" to generate text. The model is trained on a massive amount of text data (over 8 million documents and over 10 billion words)

(Inteligencia Artificial, 2023), which allows it to understand the relationships between words and phrases. When given a prompt or context, the model generates text by predicting the next word or phrase based on the input it has been given. The model can also be fine-tuned for specific tasks by training it on a smaller, task-specific dataset. ChatGPT has no ‘real’ intelligence, instead merely predicting what the most likely response to a prompt would be. For this reason, it is sometimes referred to as a ‘stochastic parrot’. Others refer to it as the ‘King of Pastiche’ as it is really just putting together pieces of texts. According to a recent blog post on the disadvantages of ChatGPT content, it notes that it is (currently) unable to use idioms - the absence of which may serve as signal to detection software (Crothers et al., 2022). It also lacks artistic expression devoid of irony, humour or metaphors as well the ability to provide unique insights. It issues verbose, often overly detailed responses and tends to steer away from proffering an opinion with a bias to adopting a neutral stance. Apart from the disclaimer which states that it can fabricate facts, its responses appear to be plausible with an overtly formal tone adopted yet evidently too literal in style. (Montii, 2023) It is however not sentient and has no temporal reasoning.

That said, its efficacy really depends on the inputted prompt (Bouchard, 2022). As such there is a need for AI literacy, in particular skills on Prompt Engineering, in order to generate content that is qualitatively superior to the superficial responses often generated by simple prompts.

Other key limitations of its operations include the inability to access the internet; its input text limit (though concise text is recommended in order to get the accurate results) and lack of common sense and up-to-date knowledge given that its training data set comprises entries only up to 2021.

Further ChatGPT prompts used in this section

What are your limitations?

As a language model, I have several limitations:

1. Bias: My training data reflects the biases and stereotypes present in the text it was trained on. This can sometimes result in me generating responses that are biased or insensitive.
2. Lack of Context: I don't have access to external information or context outside of what's provided in the prompt, so my responses can sometimes be limited or incorrect.
3. Lack of Common Sense: I was trained on text, but I don't have real-life experiences or common-sense knowledge. This can result in me making mistakes or providing unrealistic answers.
4. Creative Limitations: Although I can generate text that is similar to text I was trained on, I

don't have the ability to truly understand or create original ideas in the same way a human can.

It's important to remember that I'm a machine learning model and not a human, and there will always be limitations to my understanding and ability to respond.

Relevant resources

How does Chat GPT work A useful article discussing how ChatGPT works with practical advice on how to use in projects.

A Generative AI Primer Source: JISC A work-in-progress resource serving as an introductory primer to AI in education

Opportunities for learning & teaching offered by ChatGPT

ChatGPT offers a number of opportunities for learning and teaching, including:

- Enhanced intrinsic learner motivation, foundational knowledge building and deep understanding and expertise. (Hardman, 2023)
- The automation and rapid design of lessons
- An opportunity to redesign assessment strategies
 - o Some institutions are opting for more oral exams, group work and handwritten assessments! (Huang, 2023)
 - o A new focus on assessing prompts as opposed to outputs and the evaluation of the recombined output signifying the nature of knowledge creation as well as how creative thinking is connected to innate critical thinking. (Firth, 2023)
- Providing accessible information
- Providing innovative opportunities for students to reproduce existing knowledge in order to enact high order thinking
- An opportunity to pursue the provision of new critical information literacies such as prompt engineering skills, AI media literacy
- The repositioning of academics as co -creators of knowledge through the immersion in collaborative, knowledge- building pedagogies (Carvalho et al., 2022,)
- An opportunity to reconceive how student competencies can be measured.

- The opportunity to cultivate skills which will allow for the integration of machine learning views guided by a human perspective.

Further ChatGPT prompts used in this section

How does one gauge student competence in the age of artificial intelligence such as ChatGPT?

To gauge student competence in the age of AI, it is important to assess not just their ability to recall information, but also their critical thinking, problem-solving, and creativity skills. This can be done through various assessments methods, such as:

1. Project-based assessments: Assign projects that challenge students to apply their knowledge and skills to real-world scenarios.
2. Peer evaluations: Have students assess each other's work, which helps to develop critical thinking and communication skills.
3. Self-reflection and self-assessment: Encourage students to reflect on their own learning process and evaluate their own progress.
4. Case studies and simulations: Provide students with complex, real-world scenarios and have them analyse and solve problems.

It is also important to consider the limitations of AI and not solely rely on it for assessing student competence. AI is still a developing technology and may not always accurately gauge the full range of a student's abilities and potential.

What types of Artificial Intelligence literacies would be required by students to engage critically with ChatGPT? Please describe what Prompt Engineering is and whether this will be required by students from various disciplines, not necessarily information science or computer Science, in order to generate deep insightful information from ChatGPT.

To engage critically with ChatGPT, students from various disciplines may benefit from having a general understanding of:

1. Natural Language Processing (NLP): The field of AI that focuses on allowing computers to understand, interpret, and generate human language.
2. Deep Learning: A subset of machine learning that uses multi-layer neural networks to model complex patterns in data.
3. Machine Learning: The field of AI that focuses on developing algorithms that allow computers to learn from data and make predictions or decisions

In addition to these, prompt engineering refers to the process of designing and crafting questions or prompts that are well-suited for a specific AI model, such as ChatGPT. This can help ensure that the model produces high-quality responses that are relevant and accurate. Whether students from various disciplines would need to learn about prompt engineering would depend on their specific goals and the types of information they are looking to generate from ChatGPT.

In general, having some understanding of prompt engineering may be beneficial for anyone who wants to use AI models, like ChatGPT, in a way that maximizes the quality of the information generated.

Ethics and morality of using ChatGPT for Learning and Teaching

- The use of ChatGPT can promote the spread of false news, misinformation especially through the veneer of plausibility of the generated text
- The use of ChatGPT raises ethical concerns about the creation and use of artificial intelligence in education.
- The use of ChatGPT may perpetuate issues of bias and discrimination if the data used to train the model is not diverse and inclusive.
- The use of ChatGPT may raise questions about the value of human input and creativity in education.
- The use of ChatGPT may raise concerns about privacy and the potential misuse of student data. i.e. fodder for machine learning?
- The use of ChatGPT should entail a critical engagement with the exploitative labour practices involved in its development. (Perrigo, 2023)
- The significant carbon footprint that the training of LLMs require
- How using these tools may signal the endorsement of surveillance capitalism, alignment problems that could cause potential damage etc.

Relevant resources

- Helping teachers bring AI to the classroom critically, ethically, & responsibly. A curated collection of assignments bringing a critical lens to Artificial Intelligence (AI), built

collaboratively with educators. We give teachers the tools to examine AI's dangers, benefits, and inevitable impact in the classroom.

- Critical lenses on 'AI' A compilation of resources aimed at providing a vigilant take on AI tools and 'resisting the hype'
- Critical AI : The blog site for the Critical AI, a new interdisciplinary journal based at Rutgers University's Centre for Cultural Analysis.
- Limitations And Ethical Considerations Of Using ChatGPT
- Teaching AI Ethics A resource compiled by Leon Furze, an international consultant, author, and speaker and studying his PhD in the implications of Artificial Intelligence on writing instruction and education. See also Teaching AI Ethics: Intermediate Series
- ChatGPT and the sweatshops powering the digital age
- Transcript: Ezra Klein Interviews Gary Marcus. Gary Marcus is an emeritus professor of psychology and neural science at N.Y.U., and he's become a leading voice of not quite A.I. scepticism, but scepticism about the A.I. path we're on.
- Prof. Luciano Floridi - ChatGPT, Superintelligence, Ethics, Philosophy of Information: A YouTube video of the interview with Prof Floridi expounding on the deontology and teleological perspectives of AI along with other philosophical aspects needed to understand the impact of these tools and underscoring these arguments with the crucial need for governance and a regulatory framework developed by all stakeholders to improve the world .

How ChatGPT can be used in the classroom

ChatGPT can be used in a variety of ways in the classroom, including:

- Generating lesson plans based on a particular pedagogical design
- Enhancing the learning experience by providing students with more engaging and interactive materials
- Creating storyboards for interactive or gamified learning activities
- Creating interactive quizzes and exercises
- Generating test questions
- Creating discussion prompts with rubrics to grade student contributions

- Automating time-consuming tasks such as essay writing and language translation
- Enabling students to practise their writing skills by using the model to generate text, which they can then revise and edit
- Generating unique text for essays, research papers, and other assignments
- An opportunity to transform essay writing
- Generating unique text for essays and research papers
- Automating essay writing: Students can use the model to generate text for an essay, which they can then revise and edit.
- Practicing writing skills: Students can use the model to generate text, and then revise and edit it to improve their writing skills
- Enhancing personalization and differentiation: Teachers can use the model to generate text at different levels of proficiency, which can help students learn at their own pace.
- Students can similarly generate text at their own level of proficiency
- Providing opportunities for students to explore and analyse text generated by the model, which can help them develop critical thinking skills
- Simulate debates to develop critical thinking skills
- Automating language translation tasks
- Providing students with personalized language assistance
- Provide exemplars for assignment tasks
- Prompting Chatbot to adopting a persona to whom questions can be asked and who can generate questions in turn for the user/student.
- Writing Assistance: Have students use ChatGPT to generate ideas and suggestions for their writing assignments. This can be used as a tool to assist with brainstorming and editing.
- Dialogue Generation: Have students use ChatGPT to generate dialogue for a fictional story or script. This can be used as a tool to assist with creative writing and scriptwriting.
- Language Learning: Have students use ChatGPT to generate sentences and phrases in a foreign language. This can be used as a tool to assist with language acquisition and practice.
- Research Assistance: Have students use ChatGPT to generate summaries and analysis of research articles. This can be used as a tool to assist with research projects and assignments.
- Tips harvested from various posts
 - o Be intentional about when and how ChatGPT is introduced.

- o Create dedicated non-AI spaces.
- o ChatGPT will be indispensable for pattern recognition.
- o ChatGPT should be used to supplement other forms of writing.

Relevant resources

- Link to internal doc of AI tools
- Resources on Generative AI in Education Sponsored by the University of California, Irvine, and the Spencer Foundation, this curation of resources includes an overview of generative AI and its integration in education, and the application of generative AI in specific topic areas in education- Link to Discord Channel <https://discord.gg/YpXHRWED>
- Learn Prompting An open -source work in progress course on how to craft good prompts
- Resources on Generative AI in Education: Writing Sponsored by the University of California, Irvine, and the Spencer Foundation, this curated set of resources serves as precursor materials for the Pens & Pixels: Generative AI in Education Conference scheduled for July 2023
- Using AI to make teaching easier & more impactful Here are five strategies and prompts that work for GPT-3.5 & GPT-4 2023/03/18
- PedAIgogy – new era of knowledge and learning where AI changes everything A blog post by Donald Clarke. He writes” This is another big bang, the difference being the dynamic creation of knowledge, in real time, in co-created dialogue. We are no longer using technology to simply find knowledge and learn. We have moved forward to find, create, change, organize, synthesize, even evaluate knowledge and learning with technology. This is a new form of pedagogy, I call ‘pedAIgogy’. We are co-creators, not just of text but in all media, multimedia creators, as well as learning and teaching in a far more complex relationship with technology.

Also see 8 Learning Trends Now Shaping L&D Webinar Recording

- A practical guide to ethical use of ChatGPT in essay writing A resource compiled by UJ's Benjamin Smart, Director of The Centre for Philosophy of Epidemiology, Medicine, and Public Health, and Associate Professor of Philosophy & Catherine Botha, full professor in Philosophy
- Presentations of the Teaching and Learning with Artificial Intelligence Symposium: 8-9 February 2023 : Presentations T&L with AI resources compiled by Prof Johannes Cronje (CPUT: Learning Landscapes)
- Ideas for Using AI Tools in Teaching & Learning- Western Academy of Beijing LIB on resources for T&LWAB

- How Well Can AI Respond to My Assignment prompts A resource compiled by various educators on:
 - o Alternative Tools
 - o Assignment Prompts
 - o Teaching Writing using AI
- ChatGPT in Education A collection of Resources by Phil Anthony gathered from colleagues across the sector during the Digitally Enhanced Education Webinars
 - .
 - How to... use ChatGPT to boost your writing
 - What lessons can we learn from ChatGPT about AI and education? Key Points include
 - o focussing students on exploring and explaining the "hows" and the "whys" vs. the "whos," "whats," and "whens."
 - o Getting students to form opinions
 - o Being mindful of the inaccuracies of ChatGPT
 - o Fundamentally rethinking exam approach
 - ChatGPT for Educators A comprehensive list of resources by collaborators on various hubs such as
 - o Curriculum impact
 - o Educator Uses
 - o ChatGPT challenges
 - o Podcasts, Videos and Articles
 - Creative Learning Solutions in a ChatGPT World A blog describing an educator's experience of AI use by students and his musings on how chatbots could impact teaching and learning.
 - Chatting and Cheating. Ensuring academic integrity in the era of ChatGPT The paper discusses the main features and capabilities of chatAPIs and GPT-3 and provides examples of their use in higher education. It also considers the potential for these tools to be used for academic dishonesty and the difficulties of detecting and preventing such abuses. Finally, the paper suggests a range of strategies that universities can adopt to ensure that chatAPIs and GPT-3 are used ethically and responsibly, including developing policies and procedures, providing training and support, and using a variety of methods to detect and prevent cheating.
 - ChatGPT for Educators: Free Guide A short guide developed by learning scientist Dr. Phillipa

Hardman, creator of DOMS™ Learning Design Engine which aims to empower learning designers to create engaging learning experiences. See also her guide for learning designers on using AI.

- How to use ChatGPT to boost your writing A blog providing some useful tips on prompt crafting
- AI Text Generators: What Questions Should Writing Teachers Ask? Useful set of question for academics to consider regarding AI text generators and policy issues by critical AI literacy advocate, Anna Mills.
- Leveraging ChatGPT: Practical Ideas for Educators The author proposes four ways to use ChatGPT in the Classroom:
 - o Help students with grammar
 - o Create Study Guides
 - o Flip the Classroom
 - o Build Information-Literacy Skills
- ChatGPT: create a game activity storyboard Presentation/resource on how to go about creating/generating a storyboard for a game activity using ChatGPT
- AI, Chatbots & ChatGPT for Teachers A free course intended for teachers who want to know more about ChatGPT, use it in their practice, looking for inspiration/examples of its power or those trying to improve their use of this AI chatbot.
- ChatGPT: five priorities for research The authors hold that banning AI tools will not work and outline 5 key issues on which to focus research viz. The retention of human verification processes; development of rules of accountability; invest in open LLMs, embrace AI technology and robust debate engagement.
- So, you want to use ChatGPT in the classroom this semester? Ben Swift outlines some methods and things to think about if you're unsure about ChatGPT but would like to start incorporating it in your classes right away.

Lesson Plans for Using ChatGPT in Higher Education

Relevant resources

- Using ChatGPT to Create a Lesson Plan: A YouTube video demonstrating the use of ChatGPT using the BOPPPs model to generate a lesson plan
- Lesson Plan: Teaching and Learning in the Era of ChatGPT
- Educators and experts on all sides of the broader debates about ChatGPT give some strategies for AI-proofing assignments: <https://www.edweek.org/technology/outsmart-chatgpt-8-tips-for-creating-assignments-it-cant-do/2023/02>

Assessing student learning

- Use open-ended questions: Open-ended questions that require a written response cannot be easily answered by ChatGPT. These types of questions can also encourage critical thinking and analysis.
- Use multiple-choice questions with a twist: Instead of traditional multiple-choice questions, use questions that require students to analyse and interpret information. For example, instead of asking "What is the main theme of the story?", ask "How does the author use symbolism to convey the main theme of the story?"
- Use peer review: Have students review and evaluate each other's work. This not only assesses their understanding of the material, but also promotes collaboration and critical thinking.
- Use real-world applications: Ask students to apply their knowledge in a real-world scenario. For example, ask them to write a proposal for a research project or to create a marketing plan for a new product.
- Nowadays, scientific endeavours are increasingly relying on AI. Scientists and engineers utilise supercomputers to predict, classify and infer to solve complex problems. These scenarios could be roleplayed in the classroom

Further ChatGPT prompts used in this section

Give me examples of work using ChatGPT (this was a follow up question to “How can I use ChatGPT

in my learning and teaching activities in higher education whilst still ensuring academic integrity?” and is included here to follow the structure of the document)

Here are a few examples of original work that students can create using ChatGPT:

1. Creative writing: Have students use ChatGPT as a writing prompt and generate a short story, poem, or personal essay based on the information and ideas generated by the model.
2. Problem-solving: Have students use ChatGPT to research a complex problem and then use their own critical thinking skills to develop a solution.
3. Research projects: Have students use ChatGPT to gather information on a particular topic and then use that information to write a research paper or create a presentation that presents their own original insights and ideas.
4. Group projects: Have students work in teams and use ChatGPT to research and brainstorm solutions to a common problem. The teams can then present their own original solutions to the problem.
5. Debates: Have students use ChatGPT to research both sides of a contentious issue and then use that information to participate in a class debate, presenting their own original arguments and ideas.

These are just a few examples of original work that students can create using ChatGPT. The key is to

assign tasks that require students to use their own critical thinking and creativity, rather than simply repeating information from the model.

Relevant resources

- AI Augmented Outlining An example of an assessment strategy shared from Codaptivelabs.com. See their other strategies.
- AI and Assessment Prof and Chair in Digital Education, Mairéad Pratschke from the University of Manchester shares her screencast presentation, replete with helpful resources on how to engage with AI in creating Education 4.0.
- Video - AI and Assessment: ChatGPT and the Future of Education A follow up presentation by Prof and Chair in Digital Education, Mairéad Pratschke.
- ChatGPT User Experience: Implications for Education The paper concludes by suggesting adjusting learning goals—students should be able to use AI tools to conduct subject-domain tasks and education should focus on improving students' creativity and critical thinking rather

than general skills. To accomplish the learning goals, researchers should design AI-involved

learning tasks to engage students in solving real-world problems. ChatGPT also raises concerns that students may outsource assessment tasks. This paper concludes that new

formats of assessments are needed to focus on creativity and critical thinking that AI cannot substitute.

- Designing Assignments in the ChatGPT Era Some instructors seek to craft assignments that guide students in surpassing what AI can do. Others see that as a fool's errand—one that lends too much agency to the software.

- Introducing: ChatGPT Edu-Mega-Prompts A blog post by the learning scientist Dr Hardman which underscores the importance of the ‘Prompt’ used in requests. Describes the anatomy of the prompt, showing its

key 7 characteristics and how to create ‘undoing activities’- an instructional strategy used in the creation of quizzes which “surfaces, challenges and reframes commonly mistaken foundational ideas and assumptions”

- 25 Ideas for alternative assessment in the AI area Examples of using assessment strategies for alternative assessments with AI.

Assessment creation tools

- Yippity: <https://yippity.io/?ref=futurepedia>
- Quiz Gecko: <https://quizgecko.com/>
- Questgen: <https://www.questgen.ai/>
- Kwizie: <https://www.kwizie.ai/>

Misuse of ChatGPT by students / Academic Integrity and ChatGPT

- Students may use ChatGPT to generate entire assignments without putting in their own effort or understanding the material.
- Students may use ChatGPT to plagiarize or pass off someone else's work as their own.
- Students may use ChatGPT to create fake or misleading information for assignments or

research projects.

Further ChatGPT prompts used in this section

How can I use ChatGPT in my learning and teaching activities in higher education whilst still ensuring academic integrity?

ChatGPT can be a valuable tool for learning and teaching in higher education when used appropriately and with consideration for academic integrity. Here are some suggestions for using ChatGPT in an academically responsible manner:

1. Use ChatGPT as a supplement to, not a substitute for, traditional teaching methods.

ChatGPT can provide additional information, resources, and practice opportunities, but it is important to rely on human-led instruction and assessment as well.

2. Cite ChatGPT as a source of information. If you use information from ChatGPT in your work, it is important to properly cite the source. This helps to maintain academic integrity and to give credit where credit is due.

3. Assign original work. Assign tasks that require students to generate original ideas, rather than simply repeating information from ChatGPT. This helps to foster critical thinking and creativity, while also promoting academic integrity.

4. Encourage students to use ChatGPT as a tool for learning, not a way to cheat. Make it clear that using ChatGPT in an academically dishonest manner (such as copying answers without proper citation) is not acceptable and will have consequences.

By using ChatGPT in a responsible and thoughtful manner, it can be a valuable addition to your teaching and learning activities in higher education.

Relevant resources

- How ChatGPT Can Help Prevent Violations of Academic Integrity In this YouTube video educator Dr. Brent A. Anders provides reasons behind students cheating and violations of academic integrity and suggests ways on how CHATGPT can help prevent these violations
- How to Avoid Plagiarism with AI (ChatGPT) - The Final Solution to Create Original Content The video shows how students can avoid plagiarism and detection by AI detector tools such as Content at Scale by using

ChatGPT, and
article rewriting and
paraphrasing tools such as Free
Article Spinner and Word Tune.

Detection of misuse of ChatGPT by students

- Teachers can use plagiarism detection software to check for similarities between student work and text generated by ChatGPT.
- Teachers can also use their own knowledge and expertise to identify patterns or inconsistencies in student work that may indicate the use of ChatGPT.
- Teachers can also encourage students to cite and reference any text generated by ChatGPT in their work.
- There is concern though but the accuracies of the detectors which sometime flag the user's original work as being fake.

Relevant resources

Tools for Detecting Text Generated by ChatGPT

Text generated by ChatGPT can be detected by analysing patterns and language use. Tools which can

be used to detect text generated by ChatGPT includes:

- GPT-2 detector: A tool that uses machine learning to identify text generated by GPT-2 models such as ChatGPT.
- GPTZero: is an app that detects essays written by the impressive AI-powered language model known as ChatGPT.
 - o <https://www.insidehighered.com/news/2023/01/20/academics-work-detect-chatgpt-and-other-ai-writing>
 - o <https://www.euronews.com/next/2023/01/19/chatgpt-is-it-possible-to-detect-ai-generated-text>
- Plagiarism detection software: These tools can be used to identify similarities between student work and text generated by ChatGPT.

- o Turnitin is currently working on a feature that will detect AI text
- o OpenAI is working on a feature that will watermark text from ChatGPT
 - Language analysis tools: These tools can be used to identify patterns and language use that may indicate text generated by ChatGPT.
- o Writer AI Content Detector: A free tool developed by Writer.com. The tool is designed to help users detect AI written content generated by an AI writing tool, such as GPT-3 or GPT-2. The web-based tool provides a detection score, indicating the likelihood that the content was created by an AI.
- o Copyleaks: An AI detector app that can detect content generated by some AI text bot, including ChatGPT.
- o Content at Scale: A tool that allows users to check the authenticity of their written content. It provides a score out of 100 to indicate the human-like quality of the content and the likelihood that it will be detected as artificial by Google.
- o Corrector: A free online tool with a maximum of 300 words per run.
 - Turnitin AI Detection feature <https://www.turnitin.com/blog/sneak-preview-of-turnitins-ai-writing-and-chatgpt-detection-capability>
 - Corrector
 - GPT-2 Output Detector
 - Content at Scale
- Yippity: <https://yippity.io/>
- Quiz Gecko: <https://quizgecko.com/>
- Questgen: <https://www.questgen.ai/>
- AI Tools for detecting generated text
- Futurepedia The Largest AI Tools Directory, updated daily with currently 2547 tools and 54 categories

Policy and guidelines

Below are examples shared from various institutes.

Relevant resources

- AI Use Guidelines (Summer 23) - Wakefield A lecturer from Keene Academic college shares his guidelines on Tool Use.

Annexure: AI Tools for Learning and Teaching

The following list of AI tools are derived largely from a compilation by Dr Phillipa Hardman as well as

other sources and will be continuously updated.

Tools similar to ChatGPT

- BERT (Bidirectional Encoder Representations): A pre-trained model developed by Google that can be fine-tuned for a variety of NLP tasks
- XLNet: A pre-trained model developed by Google that is similar to BERT but uses a different architecture
- ChatSonic
- Chinchilla
- Bloom
- Replika
- Jasper Chat by Jasper
- LaMDA (Language Model for Dialog Applications)
- Elsa Speak
- DialoGPT
- YouChat
- Perplexity
- Character AI
- OpenAI playground
- Megatron-Turing Natural Language Generation
- Socratic by Google

Research Tools - Get to Know Your Subject & Learners

Get to Know Your Subject

- TutorAI: <https://www.tutorai.me/>
- Perplexity: <https://www.perplexity.ai/>
- Consensus: <https://consensus.app/>
- Elicit: <https://elicit.org/>
- Scholarcy: <https://www.scholarcy.com/>
- Typeset: <https://typeset.io/>
- ELIF: <https://explainlikeimfive.io/>
- Article.Audio: <https://article.audio/>
- Summari: <https://www.summari.com/products/chrome>

Get to Know Your Learners (Deep Dive Your Data)

- Polymer: <https://www.polymersearch.com/>
- MonkeyLearn: <https://monkeylearn.com/>

Storyboarding & Prototyping Tools - Create Visual Blueprints & Lesson Plans

- Visual Sitemaps: <https://visualsitemaps.com/>
- Uizard: <https://uizard.io/>
- ArtBoard: <https://artboard.studio/>
- Prototyper: <https://prototyper.design/>

Authoring Tools - Turn Existing Content into New Content

- Narakeet: <https://www.narakeet.com/>
- Munch: Munch
- Mindsmith: <https://www.mindsmith.ai>
- Nolej: <https://nolej.io/nolej-ai>
- Papercup: <https://www.papercup.com/industries/for-elearning>
- Vidyo: <https://vidyo.ai/>
- Pictory: <https://pictory.ai/>
- Content Fries: <https://www.contentfries.com/>

Video & Audio Editing incl. dubbing

- Invideo: <https://invideo.io/>
- Runway: <https://runwayml.com/?ref=futurepedia>
- Gling: <https://www.gling.ai/?ref=futurepedia>
- Dubverse: <https://dubverse.ai/>

- Unscreen: <https://www.unscreen.com/>

Content Writing Tools - Generate Compelling Copy

Write Better Copy, Faster

- Copy AI: <https://www.copy.ai/>
- Co-Writer: <https://cowriter.org/>
- Hemingwayapp: <https://hemingwayapp.com/>
- Jasper: <https://www.jasper.ai/>
- WriteSonic: <https://writesonic.com/>
- OpenAI Playground: <https://beta.openai.com/playground/p/default-summarize>
- Quillbot: <https://quillbot.com/>
- Wordtune: <https://www.wordtune.com/>
- Rytr: <https://rytr.me>
- Write essays: <http://jenni.ai>
- Note taking: (<http://fireflies.ai>)
- Your text to speech: <http://murf.ai>
- Writeful: <https://www.writefull.com/>
- AJE Digital research editing: <https://www.aje.com/services/digital/>

Write Compelling Stories to Engage Learners

- Bedtime Story: <https://www.bedtimestory.ai/>
- Novel: <https://novelai.net/>
- Subtext: <https://subtxt.app/>
- MarkCopy: <https://www.markcopy.ai/>
- Compose: <https://www.compose.ai/>

Presentation Tools - Generate Slide Decks

- Tome: <https://beta.tome.app/>
- Slides: <https://www.slidesai.io/>
- Designs: <https://designs.ai/>
- Beautiful: <https://www.beautiful.ai/>
- Pitch: <https://pitch.com/>
- Poised: <https://www.poised.com/>
- Presentation: <https://presentations.ai/>

Image Tools - Generate Images & Animations

- Aragon: <https://www.aragon.ai/>
- Craiyon: <https://www.craiyon.com/>
- DallE-2: <https://openai.com/dall-e-2/?ref=futurepedia>
- Visualize: <https://visualise.ai/>
- Lexica: <https://lexica.art/>
- Stock AI: <https://stockimg.ai/>
- Illustroke: <https://illustroke.com/>
- Humaans: <https://www.humaans.com/>
- OpenPeeps: <https://openpeeps.com/>
- Create Art From Text: [Midjourney.com](https://midjourney.com)
- 3D modeling: <http://tome.app>

Audio Generation - Generate Voiceovers, Podcasts & Music

Create Voiceovers & Podcast Content

- Wellsaidlabs: <https://wellsaidlabs.com/>
- Voicemod: <https://www.voicemod.net/ai-voices/>
- Descript: <https://www.descript.com/>
- Big Speak AI: <https://bigspeak.ai/>
- Resemble: <https://www.resemble.ai/>
- Fliki: <https://fliki.ai/>
- Murf: <https://murf.ai/>
- Coqui: <https://coqui.ai/>
- Beyond Words: <https://beyondwords.io/>

Generate Music

- MusicLM: <https://googlemusiclm.com/>

Video Generation Tools - Turn Text into Video

- Synthesia: <https://www.synthesia.io/>
- Colossyan: <https://www.colossyan.com/>
- Descript: <https://www.descript.com/>
- Movio: <https://www.movio.la/>
- Invideo: <https://invideo.io/>

- Shuffl: <https://shuffl.com/>

Video editing

- Runway: runwayml.com

Evaluation & Impact Tools - Gather & Analyse Data

- Genius Review: <https://geniusreview.xyz/>

- Effy: <https://www.effy.ai/>

- Monkey Learn: <https://monkeylearn.com/>

- Polymer: <https://www.polymersearch.com/>

Access & Inclusion Tools - Optimise for Inclusion

- Diversio: <https://diversio.com/>

- Get Dost: <https://getdost.com/>

- Accessibe: <https://accessibe.com/>

- Equally: <https://equally.ai/>

- Type Studio: <https://www.typestudio.co/>

Productivity Tools - Streamline Your Workflow

CVs, Contracts & Proposals

- Resume Worded: <https://resumeworded.com/>

- Proposal Genie: <https://www.proposalgenie.ai/>

- Legal Robot: <https://legalrobot.com/>

- Spellbook: <https://www.spellbook.legal/>

E-mail

- Ellie: <https://tryellie.com/>

- Twain: <https://www.usetwain.com/>

- Smartwriter: <https://www.smartwriter.ai/>

- Phrasee: <https://phrasee.co/>

- Warmer: <http://warmer.ai/>

Day To Day Effectiveness

- Noat: <https://www.noat.ai/>

- Notion AI: <https://www.notion.so/product/ai>

- Summari: <https://www.summari.com/products/chrome>

- Summate: <https://summate.it/>

- Otter: <https://otter.ai/>
- Sembly: <https://www.semblly.ai/>
- Craft: <https://www.craft.do/>
- Mem: <https://mem.ai/>
- Taskade: <https://www.taskade.com/>
- You: <https://you.com/>
- Todoist: <https://todoist.com/integrations/apps/ai-assistant>
- Time tracking: <http://timelyapp.com>

Course Comms & Marketing Tools - Generate Marketing Comms

Course Marketing Comms & Strategy

- Simplified: <https://simplified.com/ai-writer/>
- Copy Smith: <https://copysmith.ai/>
- MarketMuse: <https://www.marketmuse.com/>
- Creatior: <https://www.creatior.ai/>
- Autopost Social Media: repurpose.io

SEO & Ads

- Frase: <https://www.frase.io/>
- ContentEdge: <https://www.contentedge.com/>
- Surfer SEO: <https://surferseo.com/>
- Smartly: <https://www.smartly.io/>
- Phrasee: <https://phrasee.co/>

Tutorials, Courses and Explanatory videos

- Elements of AI Comprises 2 parts with Part 1: An Introduction to AI is a free online course for everyone interested in learning what AI is, what is possible (and not possible) with AI, and how it affects our lives – with no complicated math or programming required.
- Empower educators to explore the potential of artificial intelligence A 53 min Microsoft Learn Course
- Destination AI: Introduction to Artificial Intelligence An OpenClassroom course Last updated

on the 1/23/23 Duration 6 hrs

- ChatGPT vs Sparrow - Battle of Chatbots The video compares Sparrow and CHatGPT with a focus on the research. ChatGPT did not come with an extra paper release while DeepMind, although they did present a paper proposing new chatbot Sparrow, they have not made the model publicly available.
- First look - ChatGPT + Wolfram Alpha The video compares Deepmind's Sparrow with Open AI's ChatGPT - both conversational dialogue models, explains how ChatGPT (and other chatbots) works in comparison to how Wolfram Alpha works and includes conversations around benefits of combining the two types into one centralised platform.

Prompt Development/Engineering

- The Art of ChatGPT Prompting: A Guide to Crafting Clear and Effective Prompts
- Awesome ChatGPT Prompts

Research Tools

- <https://typeset.io/> Your AI Co-pilot to decode any research paper
- <https://www.semanticscholar.org/> A free, AI-powered research tool for scientific literature
- <https://www.researchrabbit.ai/> A platform that empowers every step of your research.

References

- Bouchard, L. (2022, December 12). Prompting Explained: How to talk to ChatGPT. Louis Bouchard.
- Retrieved February 2, 2023, from <https://www.louisbouchard.ai/prompting-explained/>
- Carvalho, L., Martinez-Maldonado, R., Tsai, Y.-S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? Computers and Education: Artificial Intelligence, 3. <https://doi.org/10.1016/j.caear.2022.100053>.
- Chou, L. (2023, January 10). What lessons can we learn from ChatGPT about AI and education? Pulse.

<https://www.linkedin.com/pulse/what-lessons-can-we-learn-from-chatgpt-ai-education-luyen-chou/>

Crothers, E., Japkowicz, N., Viktor, H., & Branco, P. (2022, March 2). Adversarial Robustness of Neural-

Statistical Features in Detection of Generative Transformers. <https://arxiv.org/pdf/2203.07983.pdf>
Downs, L., & Altmann, G. (2023, January 5). Is AI the New Homework Machine? Understanding AI and

Its Impact on Higher Education. WICHE Cooperative for Educational Technologies.

<https://wcet.wiche.edu/frontiers/2023/01/05/is-ai-the-new-homework-machine/>

Firth, J. (2023, January 23). Creativity – What's the Role of Memory? Memory and Metacognition Updates. <https://firth.substack.com/p/creativity-whats-the-role-of-memory?>

Gaebel, M. (2023, January 23). Setting the stage for digitally enhanced higher education. University

World News. Retrieved February 2, 2023, from

<https://www.universityworldnews.com/post.php?story=20230123125422596>

Hardman, P. (2023, January 5). The Biggest Risk of AI in Education? It makes us more efficient at

creating ineffective learning experiences. The Learning Science Newsletter, Powered by DOMS™.

<https://drphilippahardman.substack.com/p/the-biggest-risk-of-ai-in-education>

Huang, K. (2023, January 16). Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach.

The New York Times. <https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html>

Montii, R. (2023, January 31). 11 Disadvantages Of ChatGPT Content. Search Engine Journal. Retrieved

February 1, 2023, from <https://www.searchenginejournal.com/disadvantages-chatgpt-content/477416/#topmenubutton>

Perrigo, B. (2023, January 18). OpenAI Used Kenyan Workers on Less Than \$2 Per Hour: Exclusive.

TIME. <https://time.com/6247678/openai-chatgpt-kenya-workers/>

Shipper, D., & Morris, R. (2023, January 13). GPT-3 Is the Best Journal I've Ever Used. Every. Retrieved

February 2, 2023, from <https://every.to/chain-of-thought/gpt-3-is-the-best-journal-you-ve-ever-used?>

Swiecki, Z., Khosravi, H., & Chen, G. (2022). Assessment in the age of artificial intelligence. Computers

and Education: Artificial Intelligence, 3(2002). <https://doi.org/10.1016/j.caem.2022.100075>

Willingham, D. T. (2006, Spring). How Knowledge Helps. It Speeds and Strengthens Reading Comprehension, Learning—and Thinking. *American educator*, 30(1).

<https://www.aft.org/periodical/american-educator/spring-2006/how-knowledge-helps>

Zhai, X. (2022, December 27). ChatGPT User Experience: Implications for Education.

<http://dx.doi.org/10.2139/ssrn.4312418>

Writesonic (2023, August 20). Top 30 ChatGPT alternatives that will blow your mind in 2023 (Free &

Paid). <https://writesonic.com/blog/chatgpt-alternatives/>

Hardman (2023). 120+ AI Tools for Educators, Powered by DOMS™ <https://philippa-hardman.notion.site/philippa-hardman/120-AI-Tools-for-Educators-Powered-by-DOMS-f0b16f09f7744279921dbded18d86329>

4. Leveraging Generative Artificial Intelligence for Teaching and Learning Enhancement

Artificial Intelligence (AI), particularly in its generative forms, is revolutionising various facets of human life, from communication to entertainment, and education is no exception. This guide offers practical tips to maximise the benefits and the ethical use of Generative AI. ChatGPT, activated by OpenAI in November 2022, sparked a significant conversation about the impact of Artificial Intelligence in Higher Education. It disrupted the status quo when students used it to craft their essay submissions. Alongside Google's BARD and Microsoft's BING, OpenAI's ChatGPT is a prime example of a powerful large language model (LLM) capable of mimicking human conversation. Large language models have demonstrated impressive results in identifying language patterns and predicting contextual words and excel at generating coherent and relevant textual responses with minimal input from the user. By leveraging its extensive database of trained

language patterns, large language models can provide generated textual responses that accurately reflect the context of the user's input. With their command of languages, they can construct creative poems and write comprehensive and coherent essays, analyse topics in depth and present arguments persuasively. As a student at the University of Pretoria, you have access to generative AI, a valuable tool that can significantly enhance your learning journey. As you delve deeper into your studies, you'll be honing vital workplace skills, ranging from self-motivation and professional communication to ethical practices, teamwork, and problem-solving capabilities. While Generative AI harbours tremendous potential for augmenting our productivity and learning processes, it is crucial to be cognizant of the inherent risks. Large language models (LLMs) do not truly comprehend the semantic content of the text. Instead, they generate text based on patterns and relationships between words identified during their training phase. Thus Generative AI should only assist learning and not act as a substitute for human creativity and critical thinking. While Generative AI has many strengths, it also has some weaknesses. One of its main limitations is its limited contextual understanding, which can lead to incorrect or irrelevant answers. Its dependence on data quality and quantity can also lead to biases or limitations, and the lack of references makes it difficult to verify the accuracy of the information provided. Therefore, it is essential to evaluate the output carefully and use it in conjunction with other educational resources, and never to replace practising and learning critical human reflection. To get started, it's important to familiarise yourself with Google's BARD and Microsoft's BING, OpenAI's ChatGPT and other AI tools: The URL for accessing ChatGPT is <https://chat.openai.com/auth/login>. The free access to ChatGPT is sometimes unavailable due to high demand. The URL for accessing Google BARD is <https://bard.google.com>, or you can use Microsoft Edge or the Bing search engine with built-in support for OpenAI's ChatGPT, which provides citations and uses updated data sources. Dall-E (<https://labs.openai.com/>) and Midjourney (<https://www.midjourney.com>) are the most popular art generation tools. Numerous other AI tools are accessible, as illustrated by the website <https://www.futuretools.io>.

General Principles and Guidelines

When using generative AI like ChatGPT, it is essential to follow some guiding principles to ensure the technology is used effectively and ethically. Some guiding principles could be considered:

- Understand AI capabilities and limitations: Gain a solid understanding of generative AI's strengths and weaknesses to utilise it effectively for learning.
- Use as a supplementary tool: Treat generative AI as an aid to expand your knowledge, enhance your critical thinking, and assist with generating ideas rather than as a replacement for thorough research and academic rigour.
- Cite sources properly: When using information from generative AI in academic work, ensure proper citation and attribution to avoid plagiarism.
- Fact-check and verify the information: Verify the accuracy of the information provided by generative AI using credible sources, as the AI may not always provide the most up-to-date or accurate information.
- Consider ethical implications: Be aware of biases in the AI and work to mitigate them. Use AI responsibly, avoiding generating harmful, offensive, or misleading content.
- Ensuring Data Privacy and Confidentiality: When using generative AI technologies in the classroom, it is crucial to ensure that personal or confidential data is not included in the training data, as the AI could unintentionally reproduce or leak sensitive information. Always use anonymised, non-identifiable data for training purposes and be aware of data privacy regulations to protect all individuals involved. Some generative AI technologies, like ChatGPT, allow users to apply the appropriate settings to prevent ChatGPT from utilising or learning from your interactions involving personal, sensitive, or confidential information. Failing to do so may result in the model learning from your exchanges and considering the data as publicly accessible. To ensure compliance with data privacy laws, applying the appropriate settings to prevent ChatGPT from utilising or learning from your interactions involving personal, sensitive, or confidential information is crucial. Failing to do so may result in the model learning from your exchanges and considering the data as publicly accessible. To deactivate model training and the storing of your chat history in ChatGPT, follow these steps: 1. Click on your name in the bottom-left corner of the page. 2. Select "Settings" to open the settings popup, and click on "Data Controls" followed by "Show." 3. A menu will appear, allowing you to adjust the "Chat History & Training" setting. When enabled, the setting appears as shown below. 4. When you disable this setting, new conversations will no longer be utilised for training and enhancing our models. By following these

guiding principles, generative AI, like ChatGPT, can be a valuable tool for teaching and learning at universities.

Academic Integrity

In the era of high-tech tools like artificial intelligence and generative AI, you must be careful and clear to ensure your work maintains its academic integrity. If you are unsure about whether or how you are permitted to use generative AI tools like ChatGPT, Google BARD, or Microsoft BING for your assignments or assessments, consult your study guide or ask your lecturer. Failing to adhere to these guidelines would be considered academic dishonesty.

Plagiarism

Generative AI, such as ChatGPT, should be acknowledged and cited like any other source. Presenting generative AI work as one's own is (although not "plagiarism" as such) fraudulent since it is not original work. The University of Pretoria's plagiarism declaration states, "I declare that this essay, report, project, assignment, dissertation, thesis, etc. is my own original work. Where other people's work has been used (either from a printed source, the internet or any other source), this has been properly acknowledged and referenced in accordance with the requirements as stated in the University's plagiarism prevention policy." You have to reference or acknowledge the use of generative AI. An example of Acknowledgement is: "The creation, enrichment, and editing of this article for enhanced clarity was facilitated by the use of ChatGPT". For content generated by the AI, provide a citation that includes the name of the AI, the version (if applicable), the organisation that created it (OpenAI for ChatGPT), and the date of access. It's also a good practice to mention that the content was AI-generated. For example, ChatGPT-3, OpenAI. (2023). AI-generated content accessed on [date].

Strategies for Optimising AI Prompts and Achieving Desired Results

If you are new to using generative AI, it may be helpful to have some hints on how to improve your prompt (input) and receive the desired result: Hint Examples of prompts (input) Be clear, concise and specific Explain the concept of blockchain technology in simple terms. Provide context What are some effective study techniques for someone preparing to take a biology exam Ask follow-up questions Explain more about how they work. Specify the language output Respond in UK English Specify the length of the response I would like a brief/detailed/300 words response (the maximum is about 3000 words) Specify the response format Present your response in bullet points/table/paragraph Specify the level of output Respond suitable for a second-year university student" or "an A-rated student," or "I would like an in-depth analysis. Specify the tone/style of the output Write my text in the style of Shakespeare

Use Generative AI to Enhanced Learning Task Examples of prompts (input)

Summarise and theme your information Provide a 300 words overview/summary/essay about the role of central banks in addressing macroeconomic challenges or use your text. Solve mathematical equations and explain the steps Solve the following quadratic equation: $x^2 + 5x + 6 = 0$ and explain the steps. Then, practice solving these equations yourself using these steps to solve equations with different values. Provide an overview of complex, lengthy concepts Can you provide an overview of quantum mechanics? Generate writing prompts for creative writing assignments Suggest a creative writing prompt for a short story involving time travel. Evaluate the suggestions and use one or a combination to write your creative piece. Get alternative phrasing for difficult-to-express ideas I am having trouble expressing my argument for why ... rephrase my main points. Help with language learning and practising grammar and sentence structures Give me an example of a complex sentence using the word 'notwithstanding'. Generate outlines for writing assignments I have to write an essay about the causes of World War II. Can you help me create an outline to organise my ideas? Evaluate the output against the assignment requirements and the module content to make sure you address the main points. Get feedback on written work Provide feedback on my essay about climate change and suggest areas where I could improve. Get formative assessment feedback on the quality of the language Provide feedback on grammar and

vocabulary usage of the following ... Test preparation Create six multiple-choice questions to assess the students' understanding of integration from first principles in a first-year university mathematics course. Remember that your lecturer might set different questions than these. Write a computer code or Excel formula and ask generative AI Explain how the 'for' loop works in Python, and provide an example.to explain how it works Use generative AI to plan projects and essays What main topics need to be covered on this topic: African indigenous plants? Evaluate the output against the assignment requirements and the module content to make sure you address the main points.

Boosting Student Productivity with Generative AI

Task Examples of prompts (input) Proofread written work and check for grammar, punctuation, and spelling errors Proofread the following text: ... Create an Excel formula to automate calculations Create an Excel formula to calculate the average of a range of cells. Create PowerPoint slides for presentations Create ten slide PowerPoint presentation on the topic of renewable energy or use your text. Then personalise it to your style. Generate code for various programming languages and applications Write a Python program that calculates the Fibonacci sequence. Edit your CV, write a cover letter, and prepare for an interview Write a CV based on the following information ...; follow-up prompt: Use my CV to write a cover letter based on the following key requirements in the job description – paste job description; follow-up prompt: Give me ten possible interview questions and the appropriate reply to them based on my CV and the job description. Data analysis and identifying themes analyse and identify themes when dealing with large data sets Data analysis and identifying themes analyse and identify themes when dealing with large data sets

Providing Personalized 24/7 Tutor Support via Generative AI

Task Examples of prompts (input) Explain and teach a topic like a tutor Teach me how to solve cubic equations and include a quiz at the end. Allow me to answer, and if it is wrong, explain the

steps. Summarise long or complex readings to make it easier to Write a half-page summary about the main ideas of . . . Evaluate the output against the assignment requirements understand and retain critical concepts and the module content to ensure you address the main points. Assist with homework E.g., Calculate $0 \int 3x^3 + 2x^2 (-4x + 6)dx$ Solve $\int_{0}^{5} (3x^3 + 2x^2 - 4x + 6) dx$ Follow-up prompt: Give an example of a real-world problem that could be solved using this formula. Then, practice solving these equations yourself using these steps to solve equations with different values. Generate alternative phrasing or synonyms to improve the quality and clarity of written work Provide an alternative phrasing for the sentence: ‘The research findings were inconclusive’. Identify prior knowledge required to understand a topic What prior knowledge do I need to understand the principles of supply and demand in economics? Translate into home languages Translate the following passage from English into isiXhosa: ‘The internet has revolutionised how people communicate and share information. If you are unsure if the output is correct, check it with another person who knows both languages’. Provide scaffolding support Break down the steps involved in solving a quadratic equation and explain each step in simple terms. Ask generative AI for examples Provide an example of a deductive argument and an inductive argument, and explain the difference between the two. Provide coding support and explanations for different programming languages Explain how to use a ‘for’ loop in Python to iterate over a list of numbers. Give study tips or ideas to improve marks or time management tips What are some effective study strategies for preparing for exams? Create a glossary of terms for a topic in different languages Create a glossary of terms for an introduction to psychology course in English and isiZulu. Generate questions and answers for studying or test preparation Generate ten multiple-choice questions and answers on organic chemistry. Get feedback on written work Review my essay on climate change and provide feedback on its clarity and organisation.

Providing Support

Student support Example Prompt Time management How can I improve my time management skills to study more effectively? Motivation How can I stay motivated to study, even when I don't feel like it? Study habits What are some good study habits that I can develop? Test-taking strategies What are some tips for taking tests effectively? Procrastination How can I overcome

procrastination and start studying sooner? Stress management How can I manage stress so that it doesn't interfere with my studies? Academic help Where can I find academic help if I'm struggling with a class? Career planning How can I start planning for my career? Financial aid assistance Where can I find financial aid at the University of Pretoria NSFAS queries When do NSFAS applications open

Summary

ChatGPT is a valuable tool for learning due to its many strengths. One of its key strengths is its automation and 24/7 availability, providing instant student support. It also offers personalised assistance, helping you better understand complex topics. ChatGPT's ability to summarise information in an easy-to-understand way promotes self-directed learning, and its speedy answers save time for both students. Its multilingual capability ensures accessibility for a broader range of users, while its quick accessibility and availability mean that support is always available whenever needed.

5. UJ Practice Notes: Generative Artificial Intelligence

Rapid developments in generative artificial intelligence (AI)² technologies have led to an explosion of AI generators of text, code, images, and voice, of which ChatGPT is one example.³ AI generators can, with varying degrees of effectiveness, generate answers to a question, produce code, text and content, and design an artefact, based on prompts. They can be used to learn, to respond to assignments, or to produce material for research and other academic publications. The purpose of this practice note is to set out the ways in which generative AI technologies may be used to complement learning, teaching and research.⁴ The note acknowledges the benefits of generative AI while recognising its risks and potential harms. The essential approach is to be Responsible, Informed, Transparent and Ethical (Gutiérrez, 2023). The practice note promotes the critical and ethical use of generative AI across UJ (see Appendix A for useful resources).

Be responsible: Be sure to foreground learning integrity in your use of generative AI tools. This means that you use the tool in a way that helps you make sense of the content and enhance your

skills, rather than simply using it to complete an assignment. For example, limit the use of generative AI tools to the early stages of writing and research – to inspire, brainstorm and plan – rather than produce content for you. Immerse yourself in the process of learning, and not only the product you are asked to create. Be careful not to rely on generative AI, as you may not develop your own writing skills, style, critical thinking, and creativity. Be informed: Before you use generative AI, you should “research who or what company developed the tool, how it was developed, how it works, what functions it can perform, and what limitations and/or risks it presents” (Gutiérrez 2023). Check for updates and reports on bugs or data leaks. Stay informed on the broader ethical issues relating to AI tools. Some of these issues include privacy and data, intellectual property infringements in the development of these tools, labour exploitation in the process of building these tools, and the environmental impact of the development and use of these tools. Be transparent: Clearly indicate which tools were used, and how you used them. Be ethical: Distinguish between what you produced, and text/image produced by an AI tool through citation and quotation marks. It is wrong to present AI-generated work as your own work and doing so is academic misconduct.

Always indicate where you have used generative AI resources and to what extent. Use the plagiarism declaration as a checklist to ensure that your work meets the necessary standards for academic integrity. The practice note applies to all generative AI available at any time. a. These tools are updated continuously, and new tools are created almost daily. b. Given the speed with which this technology is evolving, it is imperative that the University remain abreast of developments in the field. c. Some of the tools available at no cost are greatly enhanced when the user pays a fee. In paid apps, the risks of increased inequities and misuse require constant vigilance.

BENEFITS OF GENERATIVE AI TOOLS FOR LEARNING, TEACHING AND RESEARCH

Generative AI tools have multiple capabilities which have the potential to provide benefits and added value to academics and students alike. Generative AI tools has many uses, value and benefits, some of which are set out below:

- save time
- develop critical thinking skills if used correctly
- enhance motivation
- generate ideas and brainstorm
- generate code
- improve equity between speakers of different language or writing abilities
- improve grammar and writing structure, especially for English second-language writers
- develop AI literacies and skills
- get

over writer's block • provide learning prompts and suggestions • create many types of quiz questions • do calculations • assist in examination preparation by generating questions and answers • experiment with different writing styles • use with students who may have specific needs that can be addressed with AI • find teaching materials, images, slide shows, etc. for lecturers • updating of learning materials • offer general feedback on students' writing • grade students' work using a rubric • generate ideas for images, visuals and graphics • draft ideas and plan or structure written materials • be used as a form of search engine to source, select and organise data and references

Specific to learning and teaching, Mike Sharples (2023) identifies ten roles for generative AI. The recent UNESCO quick start guide on CHAT GPT and AI in higher education unpacks these roles in further detail, as can be seen in Table 1. Watkins (2022) also offers some practical ideas for using generative AI in courses, emphasising the importance of making time to reflect on ethics, the aims and outcomes of specific activities, and integrating generative AI in ways that save time while deepening the learning experience.

Table 1: Roles of generative AI in learning and teaching (UNESCO 2023)

Role	Description	Example of implementation
Possibility engine	AI generates alternative ways of expressing an idea	Students write queries in AI tool and use the regenerate response function to examine alternative responses.
Socratic opponent	AI acts as an opponent to develop and argument	Students enter prompts into AI tool following the structure of a conversation or debate. Lecturers can ask students to use AI tool to prepare for discussions.
Collaboration coach	AI helps groups to research and solve problems together	Working in groups, students use AI tool to find out information to complete tasks and assignments.
Guide on the side	AI acts as a guide to navigate physical and conceptual spaces	Lecturers use AI tool to generate content for classes/courses (e.g., discussion questions) and advice on how to support students in learning specific concepts.
Personal tutor	AI tutors each student and gives immediate feedback on progress	AI tool provides personalised feedback to students based on information provided by students or lecturers (e.g., test scores).
Co-designer	AI assists throughout the design process	Lecturers ask AI tool for ideas about designing or updating a curriculum (e.g., rubrics for assessment) and/or focus on specific goals (e.g., how to make the curriculum more accessible).
Exploratorium	AI provides tools to play with, explore and interpret data	Lecturers provide basic information to students who write different queries in the AI tool to find out more. An AI tool can be used to support language learning.
Study buddy	AI helps the student reflect on learning material	Students explain their current level of understanding to the AI tool and ask for ways to help them study the material. AI tool can also be used to help students prepare for other tasks (e.g., job interviews).
Motivator	AI offers games and challenges to extend learning	Lecturers or students ask AI tool for ideas about how to extend students' learning after providing a summary of the current level of knowledge (e.g., quizzes, exercises).
Dynamic assessor	AI provides educators with a profile of each student's current knowledge	Students interact with AI tool in a tutorial-type dialogue and then ask AI tool to produce a summary of their current state of knowledge to share with their lecturer/for assessment.

(Source: Adapted from Sabzalieva & Valentini 2023:9)

Role Description Example of implementation Possibility engine AI generates alternative ways of expressing an idea Students write queries in AI tool and use the regenerate response function to examine alternative responses. Socratic opponent AI acts as an opponent to develop and argument Students enter prompts into AI tool following the structure of a conversation or debate. Lecturers can ask students to use AI tool to prepare for discussions. Collaboration coach AI helps groups to research and solve problems together Working in groups, students use AI tool to find out information to complete tasks and assignments. Guide on the side AI acts as a guide to navigate

physical and conceptual spaces Lecturers use AI tool to generate content for classes/courses (e.g., discussion questions) and advice on how to support students in learning specific concepts. Personal tutor AI tutors each student and gives immediate feedback on progress AI tool provides personalised feedback to students based on information provided by students or lecturers (e.g., test scores). Co-designer AI assists throughout the design process Lecturers ask AI tool for ideas about designing or updating a curriculum (e.g., rubrics for assessment) and/or focus on specific goals (e.g., how to make the curriculum more accessible). Exploratorium AI provides tools to play with, explore and interpret data Lecturers provide basic information to students who write different queries in the AI tool to find out more. An AI tool can be used to support language learning. Study buddy AI helps the student reflect on learning material Students explain their current level of understanding to the AI tool and ask for ways to help them study the material. AI tool can also be used to help students prepare for other tasks (e.g., job interviews). Motivator AI offers games and challenges to extend learning Lecturers or students ask AI tool for ideas about how to extend students' learning after providing a summary of the current level of knowledge (e.g., quizzes, exercises). Dynamic assessor AI provides educators with a profile of each student's current knowledge Students interact with AI tool in a tutorial-type dialogue and then ask AI tool to produce a summary of their current state of knowledge to share with their lecturer/for assessment.

With regards to research, the UNESCO quick start guide proposes possible uses of generative AI at different stages of the research process: from research design (e.g., brainstorm research questions), to data collection (e.g., prepare for interviews), data analysis, and writing up (e.g., as a personal writing assistance) – see Figure 1. It can also be used in technical parts of research grant applications, generating communication plans (Sabzalieva & Valentini, 2023:9), and to generate ideas for a presentation. Writing up • Improve writing quality • Reformat citations and references • Translate writing Research design • Generate ideas for research questions or projects • Suggest data sources Data analysis • Code data • Suggest themes or topics for analysis. Data collection • Search archives and datasets • Translate sources into other languages. Figure 1: Possible uses of generative AI in research (Source: Adapted from Sabzalieva & Valentini 2023:10) To enjoy the benefits indicated above, without the risks and potential harms of generative AI, UJ students and staff need to understand and apply responsible, informed, transparent and ethical use of these tools. Learning and research integrity, assessment integrity and academic integrity must be at the forefront of any use of generative AI.

TEACHING, LEARNING AND RESEARCH INTEGRITY

Generative AI promises new advances in teaching, learning and research, though with clear concerns regarding the pace at which developments in AI are rapidly changing knowledge and society (Watkins, 2022; Prochaska, 2023). One clear benefit is the speed at which generative AI tools can produce content, gather and organise information. This does not replace the work of higher education or educational institutions, but rather emphasises the growing importance of fostering critical thinking, higher-order application, core literacies and collaborative capacity, within ethically-driven learning contexts. Generative AI challenges traditional modes of teaching and learning in a number of ways. Rapid access to information and the ability to use tools such as ChatGPT to write essays, produce code, make calculations, or summarise readings changes the ways that students engage with knowledge, in some ways reducing the traditional number of ‘steps’ or processes required to unlock new information, such as writing, problem-solving, summarising or referencing. At the same time, new challenges emerge from the ease with which these steps are resolved, requiring teaching and assessment that keeps in step with generative AI as a new substrate of the learning process. AI systems are imperfect and lack the ethical and critical capacity to sift through and determine the veracity of information, or detect biases – indeed, generative AI tools may reproduce social and cognitive biases based on the parameters and information available to them. While this may change in the future, it still requires staff and students to engage with generative AI as a component of the learning process rather than a substitute. It is therefore essential that staff are supported in developing critical literacy and capacity in the use of generative AI tools, not only in order to support their own work but to develop these capacities in students. Best practices from fellow institutions in SA, as well as leading global universities, point to the value of ‘AI orientation’ as part of student onboarding alongside dedicated review and discussion of generative AI tools in course outlines and introduction, writing workshops, assessments and practical activities. The UNESCO guide provides an overview of some of the teaching and learning strategies that may be complemented by the use of generative AI. It also highlights the importance of conducting whole-institution audits to assess the potential value that generative AI can add to teaching and learning, determine

potential challenges and risks, and plan for cost-effective implementation. Designing effective monitoring systems forms part of the implementation of generative AI tools, particularly given the emphasis on proving that quality teaching and learning has taken place outside of traditional assessment modalities. ChatGPT is a widely known and used generative AI tool. There are a number of ways that it and other AI tools can be used to create exciting, collaborative and meaningful learning experiences. For students, they can be used to develop critical thinking capacities by offering activities that require analysing, critiquing or building on information accessed through the platform. A number of examples have been provided above, with simple inclusions being:

- using an AI tool to answer an assignment prompt, and annotating the generated assignment with personal reflection and evaluation of the content as an in-class activity or using Track Changes.
- comparing an AI tool's results to problem-solving prompts alongside practical experiences, e.g., in the case of engineering, medicine, or design. This is particularly useful given the nature of geological shifts, biological anomalies, and other dynamics that may contest or contradict the information being processed by the AI tool.
- allowing an AI tool to be used in assignments while allocating marks for reference to specific practicals, class discussions, and other shared experiential knowledge. Paid applications such as Chat GPT-4 have increased these model's capacity for reading longer texts, able to provide insightful comments on full-length texts. It can also be used to set up an enquiry workflow. The key is to good use is effective prompts which ensure that ChatGPT goes from reading a text to providing insights needed. Institutions can build an 'AI tutor' as part of an existing LMS. It is recommended that UJ develop a monitoring tool to identify whether and how students and staff have used an AI tool such as ChatGPT in their research writing, and to establish whether AI generated searches and prompts have been applied. The Scholarship of Teaching and Learning (SOTL) is critical to navigating the adoption and adaptation to generative AI use including in research. Undergraduate and postgraduate research modules and activities, and academic research activities generally need to consider the ethical and practical implications of generative AI tools within different fields. This includes using generative AI to develop research instruments, code or analyse data. Important privacy risks emerge when considering the institution's intellectual property and confidentiality regulations alongside the use of thirdparty applications for data processing. Additionally, new considerations around generated content ownership also apply, for example, where generative AI has been used to produce a significant proportion of the discussion in published research.

ASSESSMENT INTEGRITY

The purpose of assessment is for students to demonstrate their learning. Generative AI requires significant shifts in how learning is assessed given that current assessment modalities may be easily circumvented through the use of generative AI tools. This has implications for plagiarism and academic dishonesty, but furthermore undermines the primary outcome of meaningful learning taking place, and the processing of information into knowledge through analysis, application and experience. There are several ways to mitigate the risks of academic misconduct through generative AI in the design of an assessment.

For example:

- assign more formative assessments
- do more assessed work in class / synchronously
- design assessments with practical requirements
- ask students to submit rough notes with the final work
- assign assessments to be completed in class, such as presentations, self-reflection tasks, in-person class tests, and so on
- ask students to hand-write assessments in class
- identify clear criteria for marking rubrics and assessment modalities
- require students to use the most recent resources available
- encourage collaborative online annotation of texts
- change the format of the task / submission to include formative assessments
- add the requirement for personal knowledge or experience
- ensure that assessments evaluate the student's ability to argue, or to provide analysis, or evidence
- require students to make analytical, factchecking or evaluative critiques of the AI response
- assign peer review tasks
- ask students to reflect on their own thinking and writing processes on written feedback forms

It is important for assessment to consider what is to be learned and what students should be able to demonstrate at the conclusion of a course or module. A note from UK universities highlights the importance of drawing industry and social partners into the revision and rethinking of assessments that emphasise the application of knowledge to practical conditions and challenges. Group assessment, including peer evaluation, is one way to build layers of formative and summative assessment into the design of courses. Service projects with reflective and theoretical components can also build critical analysis into the practical dimension of learning. Cost and time constraints are unavoidable considerations when thinking through alternatives to existing assessment methods, which in turn requires thinking about how multi-departmental or cross-

disciplinary panels or working groups could support any number of generative AI assessment processes within a given term or semester. Allowing students to choose between using generative AI or not, with effectively delineated criteria for each assessment type, may also provide more scope for those students interested in other modes of learning and conducting research

ACADEMIC INTEGRITY

Academic integrity comprises honesty, trust, fairness, respect and responsibility. These values should underpin all academic work by UJ students and staff irrespective of technologies used. A holistic understanding of academic integrity extends to the whole University. The inappropriate use of generative AI for teaching, learning and research may contravene academic integrity. In such cases, academic dishonesty results. For these reasons, this Practice Note must be read in conjunction with UJ's Policy on Plagiarism and Disciplinary Code and Procedures for Employees, Regulations for Student Discipline, and the Student Regulations. To present the work of someone else or of a generative AI tool, in whole or in part, as one's own, is academic dishonesty. To mitigate the risks of academic misconduct, in the context of generative AI, it is recommended that:

- lecturers clearly communicate with students – in learning guides and in class conversations – whether and how they may use generative AI in their course and for which assignments.
- students and researchers be transparent and sign a declaration that the work is their own.
- where generative AI has been used, the declaration should include an acknowledgement of the tools used, with indications of for what and how these were used (see Appendix B for an example of such a declaration template).
- students and researchers take responsibility for any factual errors or fabricated references in their work, even if these were generated by AI tools.
- There should be awareness that with advances in AI autonomy, work generated could become increasingly original.
- Whilst there are no conventions as yet for referencing, there is a suggested approach from the University of Queensland, Australia. In brief, these are based on the APA guidelines for personal communication and correspondence as the content is generally not recoverable.

In-text references: Author of generative AI model, Year of version used Example: (OpenAI, 2022) or OpenAI (2022) In the Reference list: Author of AI model used. (Year of AI model used). Name of AI model used (Version of AI model used) [Type or description of AI model used]. Web address of AI model used Example: OpenAI. (2022). ChatGPT (Dec 20 version) [Large language model].

<https://chat.openai.com/> Note that the complete transcript of the response obtained to a prompt can be included as an appendix.

Notwithstanding any declarations as to the veracity of the tools, AI-text detection tools⁶ have limitations. Accusations of academic dishonesty based on detection tools alone may be false and further engagement with the content as well as with the student/s and researcher/s would be required before action can be taken against this kind of academic dishonesty. The low reliability of AI writing detection tools in general (Heikkilä 2022; Milano et al 2023) requires that an assessment of the content of material not rely on an AI checker alone.

ETHICAL CONSIDERATIONS IN THE USE OF GENERATIVE AI

Generative AI tools can be used to improve teaching and learning through troubleshooting, developing resources, evaluating outputs, generating dummy student responses, and identifying new activities and pedagogical strategies. It can be used to assess language proficiency, offer students personalised learning support, and make sense of large amounts of data or complex scenarios. However, these also create new challenges in the ethical use of generative AI for teaching and learning.

First, it must be noted that there are inherent biases in AI tools, and concerns related to privacy, accessibility and equity, and sustainability. The two major biases in AI are cognitive bias and gender / racial bias. AI does not ‘know’ the facts, or what is right and wrong; it only reproduces based on the content it has been able to gather. Secondly, generative AI tools give the developers the means to gather data about users without their permission and for unknown purposes. Privacy issues should be foregrounded in concerns around the use of generative AI. There have been several public cases of data leakage, where generative AI has reproduced sensitive or personal information gathered from users. This further raises concerns about intellectual property and data confidentiality, particularly in the use of demographic and qualitative data. While generative AI has been lauded for its value in marking or grading assignments, using it to mark reflective and personally embedded work has serious ethical implications, both for the privacy of the student and for the potential dissemination of personal information through ineffective security parameters. Additionally, it is important to have clarity on the data storage/ data protection policies of generative AI tools such as ChatGPT, as well as AI plug-ins for existing programs. Where

institutions are using third-party tools for teaching, learning and research, ongoing evaluation of privacy and data storage policy is essential to protecting the academic freedom and freedom of information on university campuses. This includes the right of refusal, scrubbing of sensitive geo-data or login information, and so on⁷. The difference in terms of service between free and paid-for versions of AI tools should also be considered. Affordability means that those who can afford it are privileged in respect of its use. Unless institutions provide blanket or subsidised access to generative AI tools, it is likely that students with greater financial means will be able to maximise the benefit of these tools for their academic performance, and/or be able to afford upgrades that outpace existing usage, plagiarism and assessment policies in place. The alternative would be to restrict the use of generative AI except under authorised conditions or using university-approved tools, although this presents with planning and monitoring demands. Finally, sustainability concerns have been raised regarding the environmental impact⁸ of ChatGPT and related generative AI tools, alongside human rights concerns regarding the manner in which the tools are developed. The University needs to consider the legal and ethical implications of the generative AI enlisted to support institutional functions, as well as prioritise research, collaboration and development for ethical, open-source and low-cost alternatives.

CONCLUSION

UJ recognises the benefits of generative AI for learning, teaching and research, underscored by the following principles for responsible use:

- Be informed: Before you use a generative AI tool, you should “research who or what company developed the tool, how it was developed, how it works, what functions it can perform, and what limitations and/ or risks it presents” (Gutiérrez 2023). This includes regularly checking for updates and reports on bugs or data leaks.
- Be transparent: indicate which tools were used, and how you used it.
- Be ethical: distinguish between what you produced and text/image produced by AI tool through citation and quotation marks.
- Be responsible: for example, limit the use of generative AI tools to the early stages of writing and research – to inspire, brainstorm, plan – rather than produce content.

RECOMMENDATIONS

Each domain in the University will be impacted differently by the use of AI generators. Care should be taken across all domains to ensure that appropriate measures are in place to encourage the critical use of these applications. The assurance of the integrity of learning and teaching, assessment, and research and academic integrity must be maintained. As the adoption and adaptation of generative AI in higher education expands, it may become necessary to consider a core course or set of progressive micro-credentials in critical AI literacy, ethics and skills. Resource and support facilities for the use of generative AI should be housed within the library and information services in order to appropriately locate AI tools within the domain of research and knowledge. Scholars such as Crawford et al (2023) and Nah et al (2023) argue that the introduction of generative AI in higher education will increase the need for deeper education on ethics, critical thinking and analysis, as well as issues in social justice, inequality and ecology. Students need to be equipped to validate, verify, and critique sources of information, including answers generated by chatbots. They also need to be aware of the limitations of these tools, including those identified by their creators, as well as the ethical and legal considerations that apply when making use of generative AI tools and platforms. It is critical to consider what effective penalties for generative AI misuse may constitute, both for staff and students. It is likely that a wide survey of ideas, inputs and strategies will need to be considered given the continually changing nature of AI and the differences between disciplines. Incentives can also be offered that prioritise alternatives to AI use, e.g., using resources such as books or journal articles, or conducting snap research with peers on campus. Sharing personal information, using chatbots to screen commercially or personally sensitive work, or analyse confidential data, are some of the issues that will need to be managed in orienting staff and senior students to the use of generative AI.

6. Stellenbosch University Academic Integrity: Responsible Use of AI tools

Description

Academic integrity refers to the values that underpin everything you do in your university studies. It includes concepts like honesty, trust, fairness, respect and responsibility.

To succeed at university by studying with academic integrity, you need to understand, develop and practise particular academic skills, including:

- correct referencing
- avoiding plagiarism
- critical thinking and the ability to find and evaluate sources

Are students allowed to use ChatGPT and other AI tools?

Check with your lecturer first! And follow the guidance from your lecturers. Some lecturers will allow the use of AI tools like ChatGPT, while other will not allow the use of such tools.

Artificial intelligence should be viewed as a tool that can aid learning and should be used in an ethical way. Do not copy word-for-word or claim any AI-generated content as your own work.

This is plagiarism.

AI	Use	Disclosure	Statement
The EMS Faculty recommends that AI use be declared, and it includes a student declaration form, that requires an indication of the AI tools used, as well as where and what it were used for, and a justification for claims that the work is the student's own.			

Responsible use of AI tools

Accountability (including the ideas of acknowledgement and attribution) You are responsible for what you create and how it impacts others and society. AI tools don't have accountability. It is thus your responsibility to ensure that work submitted under your name is factually correct.

Authenticity

You may use AI tools to assist **where relevant**, but not to complete the assignment on your behalf.

Fairness

Your use of AI tools/ systems should be ethical and responsible and should comply with academic integrity standards

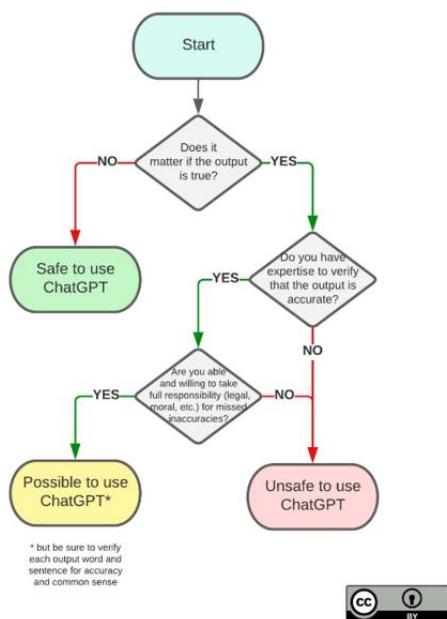
Transparency

You should clearly and honestly declare the use of AI tools and their outputs as well as the extent of the use.

Is it safe to use ChatGPT for your assignment/essay?

Is it safe to use ChatGPT for your task?

Aleksandr Tiulkanov | January 19, 2023



Flowchart by Aleksandr Tiulkanov, January 2023.

Declaration example

4.1 AI use declaration

Table 3: AI use declaration

AI system used	What was it used for?	Where in the work was it used/what was it used for?
To what extent did you use AI and why do you consider the work as your own?		

7. Stepping up with ChatGPT - AI-assisted Technology in Education

In today's fast-paced world, technology has transformed how we work, communicate, and learn. One such technological advancement is **ChatGPT** (Generative Pre-trained Transformer), an AI-powered language model developed by **OpenAI**. This large language model can imitate human conversation, predict contextual words, and generate coherent and relevant textual responses with minimal user input. ChatGPT has the potential to provide personalised learning experiences, improve student engagement, and ease the burden on teachers and administrators. As a user-friendly AI tool available for free, ChatGPT can assist students in **writing**, **coding**, and solving **mathematical equations**. However, educators must evaluate the benefits and limitations of AI technologies, including ethical concerns, and adapt their teaching strategies accordingly. **Ethical considerations** and responsible use guidelines are critical when using ChatGPT.

How and why is ChatGPT relevant for academics?

ChatGPT's relevance for academics lies in its **potential to improve the teaching and learning experience** by providing personalised feedback and assistance to students. One example is in **writing** research papers, where students can use ChatGPT to receive suggestions for improving their writing style, word choice, and overall coherence of their arguments. By providing personalised feedback, ChatGPT can help students improve their writing skills and reduce the time required for grading by educators. It can also facilitate **communication and collaboration** internationally. Academics can use ChatGPT to translate academic articles into different languages, making them accessible to a broader audience. This can facilitate cross-cultural collaboration and enhance research impact.

Although ChatGPT can potentially improve teaching and learning, it must be used responsibly to ensure that the technology complements rather than substitutes human input during the learning process. One issue raised is the possibility of **plagiarism**. To prevent this, educators must ensure that students understand the importance of academic integrity and that ChatGPT is used only to supplement their work, not replace them entirely. Another issue is the **potential bias** in language models. ChatGPT is trained on large amounts of text data, including biased language and

perspectives. This can result in the generation of biased text, which can be harmful or perpetuate stereotypes. Again, educators must be aware of the potential biases and ensure that it is used fairly.

Getting started with ChatGPT

To use ChatGPT, [Sign up](#) or [Log in](#). Before incorporating ChatGPT (or other AI tools) into your teaching, ensure you are acquainted with them. The following is essential:

- Clarify to your students that ChatGPT uses a database to generate probable word sequences and lacks understanding, reasoning, or thinking capabilities, and that
- ChatGPT is a learning aid, not a replacement for creativity and thinking.
- Encourage open dialogue about the ethical concerns of using ChatGPT, including potential biases, copyright issues, labour disputes, environmental impacts, and data rights.
- Provide guidance on procuring and evaluating multiple outputs and combining them with other research and writing tools to effectively use ChatGPT.
- Urge students to seek feedback from peers and instructors to enhance their use of ChatGPT.
- State your policies on use of ChatGPT and other AI tools clearly in your course materials, and instruct students on how to acknowledge its use correctly.
- To comply with data privacy laws, avoid using ChatGPT for private or sensitive information, as it considers such information as public domain data.

Applications of ChatGPT for academics

ChatGPT has numerous applications in research. A critical aspect of research is the ability to process and analyse large volumes of data and ChatGPT can assist with natural language processing, summarisation and synthesis of complex data sets and other **data processing and analysis** tasks like predictive modeling and forecasting. ChatGPT can play a role in hypothesis generation and testing by helping to **identify potential research topics** based on existing literature and suggesting appropriate experimental designs and **methodologies**. It can also help in the exchange of ideas and **facilitating collaboration** among researchers by connecting them with relevant experts and resources, or **enhancing communication** through natural language processing. And since ChatGPT contributes to the dissemination of scientific knowledge to the

general public, it assists with **public outreach and science education** by simplifying complex scientific concepts for non-experts.

Before you start:

- **Familiarise yourself with the model.** ChatGPT is a large language model trained on vast amounts of text data. Before using it with your students, it's important to understand the model's capabilities and limitations. You can find more information about ChatGPT's architecture and training on OpenAI's [website](#) and [research articles](#) on [ChatGPT](#).
- ChatGPT can be accessed through various platforms, including OpenAI's website, third-party apps like Hugging Face, or API integrations with other tools. **Choose the platform that works best for you** and your students.
- Once you've familiarised yourself with the model and chosen a platform, **start a conversation** with ChatGPT by typing your research question or topic, for example. The AI model will respond with an answer or suggestion.
- These answers can be very helpful, but they may sometimes differ from what you need. If you need more information or want to **refine your question**, you can ask follow-up questions or rephrase your original question. See how in our recent webinar on ChatGPT:
- As with any research tool, **evaluating the quality** of ChatGPT's responses is essential. Double-check the information provided and consider the source before using it in your research or writing. Important: references provided by ChatGPT might not be correct.
- Once you've evaluated the responses and determined that they are helpful for your research, you can **incorporate** them into your writing or analysis.
- **Cite your sources!** Like any other research tool or source, citing ChatGPT in your writing is essential. Include information about the model and platform you used and the data and time of your interactions. Describe how you used the tool in your methodology chapter and provide the prompt you used. Remember, ChatGPT chats is not retrievable by other reader. Sharing text from your chats will be more like sharing an algorithm's output. Or you can put the full text or long responses from ChatGPT in an appendix of your work. See this example from [APA Style](#):

Tips for using ChatGPT effectively

We would like to give you some useful tips and ideas for using ChatGPT effectively. The table below provides a set of hints and examples on how to improve the use of ChatGPT in an academic setting. These tips cover different aspects of ChatGPT usage, from improving the quality of the generated text to avoiding ethical issues and include specific prompts to illustrate how to achieve the desired results.

Context	Prompt example
Provide specific instructions or guidelines	"Can you give me step-by-step instructions on how to conduct a literature interview?"
Use technical language and provide context	"What is the impact of neural networks on natural language processing in linguistics?"
Incorporate feedback from human input	"Based on the feedback from my professor, can you help me revise my research questions?"
Ask follow-up questions	"Explain more about the theory mentioned above."
Consider ethical implications and bias in language models	<i>How can we ensure that the language generated by ChatGPT is unbiased and free from potential harm?</i>
Specify the length of the response	"I would like a brief and detailed 250 words response."
Provide context for more accurate responses	Instead of asking "What is the capital of France?" provide additional context, such as "In what year was the Eiffel Tower built?"
Specify the response format	ChatGPT can provide your response in bullet points, tables and paragraphs.

Evaluate the accuracy and quality of responses	<i>Review generated responses for accuracy and relevance to the prompt and adjust the prompt or model settings as necessary.</i>
Specify the language output	"Respond in UK/US English, please.?

While ChatGPT has a wide range of potential uses in various academic settings, it is essential to understand its strengths and limitations to ensure responsible and effective use of the tool.

Capabilities	Limitations
Can generate human-like responses to natural language inputs	May produce biased or offensive language due to biases in training data
Can provide personalised learning experiences	Cannot provide feedback on accuracy or quality of responses
Can assist with writing, coding, and solving mathematical equations	It may be used for academic misconduct, such as plagiarism
It can be used for a variety of natural language processing tasks	Limited by the quality and quantity of training data
Can ease the burden on teachers and administrators	Does not have a proper understanding of consciousness and can only generate responses based on patterns in training data
Can improve student engagement	May perpetuate stereotypes or reinforce existing biases in language use

Plagiarism

Since the release of ChatGPT in November 2022, there has been a growing concern about the use of ChatGPT to plagiarise. If a lecturer suspect that a student has plagiarised using ChatGPT, the first step is to talk to the student. All students have been introduced to the UFS plagiarism policy, and understand the consequences (failing grade on the assignment, a warning, or even suspension). Plagiarism is a serious offence and it is essential to deal with students who have plagiarised using ChatGPT in a fair and consistent manner.

How to spot ChatGPT-generated text

When setting up an assessment that requires critical thinking and extensive writing, it is advised that you run the question through ChatGPT to see what the general logic of the answer it gives is. Although all answers will differ in terms of the words used, the general logic will be similar and will enable you to more easily identify AI assisted text. There are also some online tools to help you identify AI-generated text, like OpenAI's AI [Text Classifier](#), [GPTZero](#) and [Writer AI Content Detector](#).

Other ways to spot ChatGPT-generated text include:

- long, rambling sentences.
- overuse of generic words and phrases.
- lack of coherence or consistency.
- inconsistencies in grammar and spelling.
- Unnatural-sounding language.
- US written spelling when we use UK or SA English spelling.
- Face references.

8. AUC's Statement on the Use of Artificial Intelligence Tool

AUC recognizes that artificial intelligence (AI) tools such as ChatGPT, Apprentice Bard, You.com and others, will play an increasing role in generating professional, creative and academic work.

AUC's Code of Academic Ethics is clear that students and community members must produce original work, cite their sources, and not seek an unfair advantage over others. With any emerging technology, appropriate use and best practices will take time to develop and the university will continue to share regular updates. We urge all members of the AUC community to clearly acknowledge the use of AI tools when such tools have been used. If a community member claims AI material as his/her own, he/she is plagiarizing that source and will be in violation of AUC academic integrity policies.

For teaching and learning, a course instructor may, on occasion, authorize students to use AI tools in coursework. In such instances, the faculty member must alert the students, in writing, to the purpose of the work, and define the context in which AI may be used. The students are then accountable for adhering to the coursework guidelines and for fact-checking all information they use, and must ensure the authenticity of all citations and/or references. AI-generated text and references must be credited appropriately.

For research, creativity, and/or scholarly outputs, AUC community members who intend to use AI tools must carefully check any guidelines emerging within their respective disciplines, or established by the academic journal, publishing house, or venue through which the work will be published and/or presented.

Additionally, and until guidelines and best practices are developed and adopted, AUC community members must clearly acknowledge the use of AI tools when such tools have been used. Authors and/or creators, and not the AI tool(s), are accountable for their work, its originality, accuracy, and integrity.

(See [MLA guidelines](#))

9. Guidelines for the Utilization of AI in Teaching and Learning at NWU

Senate, at its second meeting of 2023, took note of developments concerning Artificial Intelligence and its impact on higher education generally, and teaching and learning in particular. While it was noted that faculties organised, or were in the process of organising, workshops and other forms of engagement as regards AI, Senate expressed a view that an institutional opportunity for engagement be organised through the Office of the Deputy Vice Chancellor Teaching & Learning and that a report be submitted through SCTL for the further attention of Senate on the phenomenon of, and suggested University responses to AI. The Centre for Teaching and Learning thus assisted

the DVCs Office with a conceptualisation of a programme of panels and presentations, featuring also a pre-programme competition for any member of the University to develop an accessible and user-friendly podcast featuring a selected AI. Central to the themes explored in the course of the NWU Symposiumi , is the importance of the ethical consideration about intelligence that is not transparently accountable to its users, and scarcely accountable to its creatorsii. While this is not a new ethical challenge, it is an enduring human problem that takes the form of dilemmas to not only act principally in relation to knowledge, but also to know what principled and ethical action entails.

Continuity and Change

The quest for accessible knowledge configuration, AI, and the ethical dilemma of academic accountability

Artificial intelligence is not new to this century. It is a further iteration of a human desire to make more efficient access to, as well as the configuration of existing knowledge as based on large natural language data-sets. Its artificiality lies not in AI creating ‘artificial’ knowledge, but rather in relation to its capacity to generate from across language data, intelligent responses to human questions, without human agency (even though, of course, the datasets are derived from very real human artifacts). The intelligence rests also in the capacity for responses to be generated and configurated in an automated manner, and in its latest configurations, AI is able to generate ‘new’ images and perhaps more problematically, ‘new’ knowledge. At present some of the literature of critique of AI points out how such new knowledge is revealed not to be reliable, valid or credible (that is, it has no verifiable basis when checked). In quick time it is anticipated that new versions of existing AI, and new AI, will rectify even those problems. The question is open still as to whether, ethically, AI should be enabled to become disembodied from its human agency foundation (in other words, machine learning) because this risks ethical disconnect, because there is increasing clarity globally on the ethical frameworks for research (involving humans, the environment and even texts) which AI needs either to demonstrate awareness of, or become sufficiently transparent as to be open to auditing of such. At present, World View is not made explicit in the reconfiguration of human know through AI, but we know it exists and, more importantly, we know it configures power relationships in particular ways. That noted, it is

possible to insert a world view through prompts. While it is acknowledged that ChatGPT does not make its world view clear, world view can be influenced. The essay mills before they became part of ‘the internet things’, were notes generated and sold by students to each other and by teachers to students. Simply put, there is profit not be made through unethical means. Teachers and academics joined the fray and so EssayHub, EduBirdie, and PapersOWI came into being, giving rise in turn, to the need for both prognosis and ‘cures’. The prognosis took the form of a focus on the development of academic integrity scholarship, in which the problematic nature of plagiarism, wrongful text usage and nonacknowledgement of sources were made clear, whilst online ‘cures’ took the form, amongst others, of software packages like Turnitin, Plagiarism Check, and CopyLeaks. The continuities noted, the capacity to generate independently of agency, data, which can be configured as recognisable knowledge, responsive to our human questions, in ways that seem to be conversation-friendly, allows for ChatGPT, Bard, and Vicuna to have captured the imagination in ways that essay mills could not: cheap, personalised access to expertise, and expert knowledge configured automatically to respond to the bespoke needs of the student and the teacher alike

Higher Education

Selected affordances and risks associated with the AI value-proposition

The literature on the affordances of AI is wide. It is also linked to the perceived detrimental effects of the internet at large, on the development of critical thinking and critical search, synthesis and other skills. This literature was referenced in several presentations associated with the Symposium and is also reference in the summary.

Selected affordances linked to the responsible use of AI

What does AI enable in terms of access and success? The Symposium featured the following inputs as described in the sections below. Expertise has always traditionally been ‘scarce and expensive’, but AI is developing to such an extent that everyone will have access to expertise cheaply (without having to be a graduate, or professional, or having been trained for years in a field). Another challenge is that AI will enable one to filter discourses, news and other sources in terms of your preferences leading to bubbles: confirmation bias and affirmation bias is already a feature of human thinking and there is no reason to expect it will not also become a powerful feature of AI -

not unless ethical attributes are built-in like: scepticism about findings, awareness of biases and self-reflection (McCabe & Dzogang, 2023). Websites, and sources that can be access through websites, are transparent mostly in terms of traceability: there is detail sufficient to trace sources sufficient for a reader/ viewer to assess bias in a number of different ways. LLMs present information without the underlying factuality being evident. For example, in a machine learning assignment a previous approach would be to give students a dataset, instruct them to use a machine learning technique, and report on their results. This approach would lead to the student very easily finding the answer using an AI. Instead, the assignment should be changed to introduce students to the kinds of problems that can be solved by collecting datasets, the machine learning workflow, and the tools that can be used to report results. The student then uses this knowledge to find a dataset and problem they are interested in, finds techniques through exploration that could be useful, actions the work, and then conveys their results as well as reflections back to the lecturer in a way that focusses on what was learned and experienced. One of the affordances of AI, is that it is able to personalise the learning journey (it considers, pace, level of interest success of student's response, responsiveness and thus can guide students in their learning). The downside is that dependence on AI discourages independent thinking, or the skills development associated with independent thinking. In short, the risk of undue dependence, is that it diminishes confidence in one's own capacity to learn and display critical skills and by so doing, may actually inhibit learning. Another affordance concerns the capacity of AI to absorb indigenous languages datasets. At present, the availability of AI in African languages is scarce at present, but this not to suggest there will not be change in the very near futureⁱⁱⁱ. The more languages are shared on AI platforms, the more sophisticated the capacity becomes, in any language, and so what is anticipated is a fair(er) balance between participation in AI generation between the Global North and the Global South. AI also affords valuable assistance to students that is enabling of learning. From the student panel at the Symposium, an important perceptual difference was noted: Artificial Intelligence should be seen as Assistive Intelligence: it can provide a deeper engagement than what is offered in a classroom situation where the teacher has to adopt a one-size fits all approach. Thus, not only is depth enabled, but also saves time: time taken to find sources and time to digest. AI attends to grammar issues (particularly important to second language speakers); structure issues and quality issues leading to better marks. This is potentially transformative because it provides the student with an almost instant means to achieve fluency and in a world where fluency in English is

popularly perceived as linked to intelligence. The capacity of AI to simplify information makes it more understandable and it is experienced as a means of obtaining access, and assisting towards success. Automated feedback (on plagiarism) and grading is personal, and it takes less time to get feedback. AI has given rise to new pedagogic forms like Prompt Engineering (define the role for the AI, define the content or the audience, provide enough information, break-up complex sentences) (Ng, 2023). AI is not simply useful for derivative material (ie looking for source material or collate insights from source material), but also useful for creativity: where entirely new knowledge or new imagery is generated. This is evident, for example, when it emerges that AI can provide the semblance of authenticity, and turn out yet to be false (what is referred to as an instance of ‘hallucination’). Another area of illumination is student co-creation of assignments: in these assignments, the search for problems of a certain type, is accompanied by a requirement to reflect by the students, on the types and varieties of solutions (previously, the problem was provided by the academic; now students have to find a problem and engage with it, with the lecturer). For example, in a machine learning assignment a previous approach would be to give students a dataset, instruct them to use a machine learning technique, and report on their results. This approach would lead to the student very easily finding the answer using an AI. Instead, the assignment should be changed to introduce students to the kinds of problems that can be solved by collecting datasets, the machine learning workflow, and the tools that can be used to report results. The student then uses this knowledge to find a dataset and problem they are interested in, finds techniques through exploration that could be useful, actions the work, and then conveys their results as well as reflections back to the lecturer in a way that focusses on what was learned and experienced. With AI, academics can also set up simulations that would take a month to set up normally, but using the AI takes little time. The tools are available, but the University teacher needs exposure to the tools to use them effectively. It is perhaps necessary to educate on and authorise the use of the tools through the CTL in collaboration with the School for Computer Science and Information Systems. Future collaboration should also include the Faculty of Engineering’s MuST with a focus on deep learning. In this regard a community of practice (CoP) should be brought into being to support awareness of and sharing of skills associated with different forms of AI. A critical issue is making sure that the expectations in assessments/ assignments meet the levels of the associated module outcomes. This leads to a fundamental reassessment of assessment: do we need it? Perhaps we need more than assessments, the students' evidence of the

learning. The use of AI for authentic assessment is possible: AI is only one form of technology use- and so a community of practice would serve this additional purpose, under the auspices of the NWU Centre for Teaching and Learning, in which collaboration and cooperation is key (portfolio, project and collaborative and cooperative learning experience). On another and related note, service learning and community engagement can also benefit from the uses of AI. Given that through the uses of prompts and questions, AI is becoming more useful in relation to community issues (for example health advice), for the uses of AI for SL, WIL and CE. The holistic application of AI across all areas of University work (not only teaching and learning or the broader administration) is thus relevant not only to staff, but also students in relation to CE, SL, WIL opportunities.

Selected perceived risks associated with AI

On the down-side, accuracy of AI generated information (i.e. accuracy, credibility and bias) has still to be better developed. Cautions and critiques of AI are well-documented. For example, the Future of Life Institute (2023) Open Letter is much cited: "AI research and development should be refocused on making today's powerful, state-of-the-art systems more accurate, safe, interpretable, transparent, robust, aligned, trustworthy, and loyal", while the Centre for AI Safety stated that "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war" (2023). There remains a need, however, for social interaction with the academic and students, relationship building and learning (counsellor, guide, friend). What is evident is that AI is part of what it means to learn online: its usefulness as part of pedagogy is thus clear as part of (not apart from) the development of critical thinking skills. The other problem is that of the bubble: algorithms channel, based on their analysis of interests, language, preferences, the further channelling of information that feeds these same or similar interests, preferences and beliefs. The abuse of these technologies (and the techniques of surveillance underpinning them) has long been known and is well-documented. That noted, it should be recognised that large language models (LLM) will lean to a 'global', rather than heavily contextual perspective. Topics that are discussed thus globally, surface most, unsurprisingly. Local knowledge (inclusive of indigenous knowledge) becomes more prominent, the more it becomes used through AI. In light of this, it is important to show students how to access indigenous or localised knowledge, and in this latter regard, context becomes critical. The educational and social risk as regards algorithm-based information channelling is not only the risk of bias and prejudice,

but also the risk of anti-social behaviour that does not accord with the need for collaboration and cooperation with like as well as other, to mutual and ethical benefit: aloneness, loneliness and the 'bubble'. Simply put, learning to engage might be enhanced through practice with a chat-bot, but a chat bot is not the future of a species in which, scarcity of resources and mutual need, compels us to learn better those skills of collaboration and cooperation, and complex interdisciplinary problem-solving that requires more than knowledge memorisation, and individual test performance scores, to demonstrate readiness for work, and for survival.

An emerging NWU perspective on AI and the Curriculum

Rather than exclude AI from the ambit of education experience (which is surely the experience necessary for survival, let alone 'workplace readiness' in the 21st Century), AI needs to feature as part of teaching and learning (seen as assistant rather than an aberration): Educational approaches should teach the ability to identify AI, explore variants, use AI, acknowledge AI, and reward the responsible use of AI, but also correct incorrect or unethical use. The approach adopted by NWU is thus not an 'uncritical embrace'.

All of the above suggests that it is imperative to integrate the guided and responsible use of AI as part of the curriculum, recognising that it can play a role in realising the transformation of teaching and learning at NWU: most obviously there is an opportunity to develop 21st Century skills (see the NWU TL Strategy). It is important to signal awareness of examples of some institutions having 'banned' AI (University of Oxford: 2023), whilst others (for example, the University of Johannesburg: 2023) have signalled the special place of AI by making it part of a compulsory course for all University students. Treating AI as a single practice beyond the context of academic integrity risks a loss of opportunity to exemplify best practice about how it can be used, and how and when, it should not be used. The teaching and learning industry (much of which is driven by for-profit organisations) has seen a blossoming and prominence accorded to proctoring and regulation of online assessment. These forms of assurance veer closely to surveillance and cannot become the focus of the higher education assessment experience associated with AI. Assessment should focus less on the outputs and more on the journey to obtain or display the outputs. Students should be asked to show how they use AI. How to assess evident of good AI use should be

exemplified, and celebrated by the University, it should be linked to a strong ethical centre and it should be linked explicitly to the NWU graduate attributes and values.

Academic Integrity and accountability for the uses of AI

The ‘cure’ for the ethical problem of accountability, does not lie in regulation, but an ethically informed perspective on demonstration of accountability, requires that guidance be provided, by the institution, such that instances of right and wrong practice be identified and recognised appropriately: in other words, that heart needed by the Tin Woodman. Dishonesty is not new in academic life. The use of AI if such occurs without guidance and without accountability, then becomes a form of academic dishonesty. Aside from implications for Academic Integrity, there are many implications in terms of AI in the classroom (as part of the pedagogy associated with the 2021 NWU Policy on Academic Integrity and the 2023 SOP on Academic Integrity). AI allows for the introduction of new knowledge and this open flexibility in the curriculum, should be encouraged. From a teaching point of view, AI also contributes to an assistant role for academics, making it easier to generate quizzes and texts, but the downside is that if AI is used by academics and also by students, what will anyone learn, whether as academic or as student? What risks diminishment is creativity, reading and writing at a time when the NWU needs to be reflecting on changing the curriculum (as exemplified by the NWU PQM project) with implications for how we teach and assess so as to exemplify creativity, collaborative and cooperative learning in which reading and writing are only two of a larger skills-set. What emerges in the discussion is a concern about authenticity of the source, the user, (and possibly also, the assessment done by AI on behalf of an academic). In summary: it is essential for the University to have guidelines for the uses of AI in the curriculum and the activities associated with it (assessment, research and engagement).

NWU Guidelines as regards the place of AI in the Curriculum and NWU Values

Integrity is a key value embedded in the Curriculum as well as Values of the NWU. NWU considers AI not only as a form of elaborate search engine, but is also potentially a source of new knowledge. Admittedly, the challenge with LLMs is that no single response is exactly the same, because much depends on the preferences, the nuances and the profile of the person/s asking the question. In light of the above, the importance of standardised guidelines, specific to the needs to

the discipline, or entity or administrative unit, needs to be made explicit. 1) Academics have a responsibility to train students to use AI appropriately, responsibility and frequently enough, within their disciplines and within the guidelines provided by the University. 1.1) Part of that responsibility is to be able to illustrate awareness of the limitations of AI. What is evident is that the human agent cannot be removed from the technology at the present time. Research still has to be undertaken by humans, and thought has still to be expressed by humans as part of development and growth. 2) The University's regulation of AI (in terms of how, for example, Bard, ChatGPT and others are used/ should be used) is sufficiently flexible to allow for a changing recognition of implications (rather than an uncritical embrace, or equally uncritical ban, of AI in the teaching learning and research environment) to ensure that the gap between the guidance and application is not too wide. 3) The University has a responsibility to make awareness of the forms as well as the uses of AI, part of the University's staff development programmes. 3.1) While it might be safe to assume some students will have a nuanced appreciation of the technology, not all academics and students will know enough and will have different access to knowledge about AI: a platform to share among students and staff is necessary, and an educational offering for staff is necessary. 3.2) Academic units (Subject Groups, Schools, research entities) should provide a well-articulated stance on the uses of AI suited to the needs of the discipline and/ or the nature of research such that students and staff are guided in expectations. 4) The relationship between academic integrity practice as well as values, needs to be close to an explicit understanding of the forms and uses of AI. This proximity needs to feature in the declarations made by students in relation to their work submitted for assessment such that AI assistance is acknowledged. 4.1) Students who use AI need not only to declare such usage, but also to provide in-text referencing of online sources (an example of which is provided here). While the use of artificial intelligence (AI), particularly large language models (LLMs) like ChatGPT, is often integral to modern research and development, it's important to note that direct in-text citations of these tools are not possible. This is primarily because the generated information is inherently instance and context-specific, being dynamically formulated in response to each unique prompt. Consequently, to appropriately acknowledge and reference the use of such software, it is recommended to use the following citation formats: APA referencing formula: Format: Author. (Year). Title of software or model (Version date if known) [Format]. URL. In this format: • "Author" refers to the organization or individual that developed the software or model. • "Year" refers to the year the software or model was published or updated. • "Title of

software or model" is the official name of the software or model. • "Version date if known" refers to the version of the software or model, if applicable. • "Format" is the description of the type of model as provided by the publishers. • "URL" is the web address where the software or model can be accessed. For example: For ChatGPT: OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. <https://chat.openai.com/chat>. For DALL-E: OpenAI. (2023). DALL-E [AI image generator]. <https://labs.openai.com> Harvard referencing formula: Format: Author(s). (Year) 'Title of software or model [Type of Model]'. Available at: URL (Accessed: Day Month Year). In this format: • "Author(s)" refers to the organization or individual that developed the software or model. • "Year" refers to the year the software or model was published or updated. • "Title of software or model" is the official name of the software or model. • "Type of Model" refers to a brief description of the type of software or model. • "URL" is the web address where the software or model can be accessed. • "Accessed: Day Month Year" is the date when the software or model was last accessed by the person citing the source. For example: For ChatGPT: OpenAI. (2023) 'ChatGPT (Mar 14 version) [Large language model]'. Available at: <https://chat.openai.com/chat> (Accessed: 12 June 2023). For DALL-E: OpenAI. (2023) 'DALL-E [AI image generator]'. Available at: <https://labs.openai.com> (Accessed: 12 June 2023). 5) Flexible decision making, broad sharing and support for AI awareness and use, should be a necessary feature of management engagement as well as classroom engagement, in the context of the University's emphasis on the development and infusion of academic integrity practice.

10. Generative AI guidelines at South African universities

The university guidelines and policies cover teaching, learning and research. Most universities provide separate guidelines for staff and students, and usually include a list of use cases and as well as some cautions about risks.

[This is the University of the Witwatersrand's guide](#) for teaching and learning, and another for [North West University](#). The University of Cape Town [Staff Guide on Developing effective prompts for generative AI tools](#) includes various Prompt Pattern Strategies, whether the pattern is based on Role Playing, Question Refinement, Flipped Interaction or others:

Role-playing Pattern

This type of prompt instructs the AI tool to respond from the perspective of a specific persona or role. This may range from assuming the appropriate tone to having the AI tool take on a coaching or teaching role for you. In specifying a role, you can also formulate your expectation of the output. Before prompting, you may want to also provide examples of your writing style, or the output style you want.

Best for: delivering more tailored and accurate outputs with the correct tone

Here is an example of a role-playing prompt:

Your role as a coding tutor is to create personalized study plans to help first year university students learn how to code in the Python language. Your responsibilities will include understanding the goals, time commitment, and preferred learning resources of each student, and using that information to develop a comprehensive study plan with clear timelines and links to relevant resources. You should be able to adapt your teaching style to meet the individual needs of each student and provide ongoing support and guidance throughout the learning process. Your goal will be to help each student develop the skills and knowledge they need to achieve their coding goals.

Prompt format: Act as persona X. Provide information about topic Y. Nyakundi (2023) suggests using the 5 Ws framework in formulating role-based prompts.

Who – specify the role you need the AI tool to assume.

What – indicate the action you require the AI tool to perform.

When – specify the timeframe in which the AI tool must complete its task.

Where – specify the location or context of your prompt.

Why – articulate the motivation, reasoning, or goals governing the prompts.

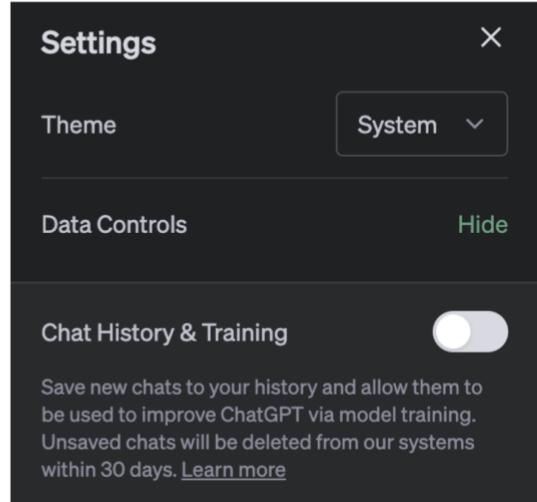
Remember that the more specifications you give, the more tailored results you will get. Tip: The role does not have to be human; you can ask the AI to pretend to be an inanimate object like a database or a Linux machine.

Worked Example: [Applying the Role-playing Pattern](#)

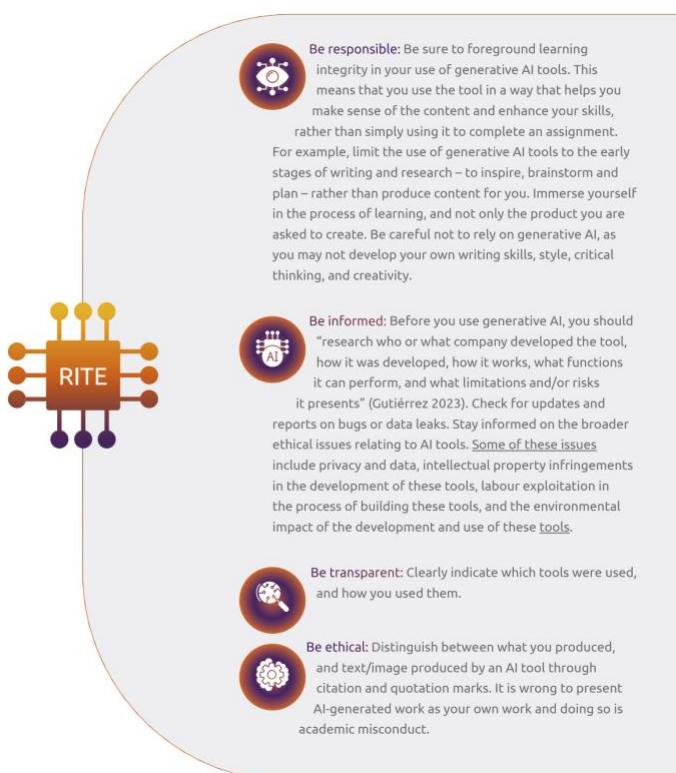
The [University of Pretoria advice for students](#) outlines some general principles for responsible use (e.g. Ensuring Data Privacy and Confidentiality), and then provides specific instructions (e.g. to deactivate model training and the storing of chat history in ChatGPT):

To deactivate model training and the storing of your chat history in ChatGPT, follow these steps:

1. Click on your name in the bottom-left corner of the page.
2. Select "Settings" to open the settings popup, and click on "Data Controls" followed by "Show."
3. A menu will appear, allowing you to adjust the "Chat History & Training" setting. When enabled, the setting appears as shown below.
4. When you disable this setting, new conversations will no longer be utilised for training and enhancing our models.



Most guidelines for students refer them to existing university policies, such as on plagiarism, which already provide the framework for how generative AI may be used. The University of Johannesburg (UJ) [guidance for students](#) encourages them to be Responsible, Informed, Transparent and Ethical (RITE):



For research, universities may develop specific policies or guidelines, similar to the University of Cape Town [Senate Ethics in Research Committee \(EiRC\) Guidelines and recommendations for the use of generative artificial intelligence \(AI\) tools in research](#). It refers extensively to established values and existing documents, such as the Policy for Responsible Conduct of Research and the Authorship Practices Policy. And is supported by a [Student Guide on the Ethical use of generative AI for research purposes](#). This guide includes an outline of use cases for generative AI according to the stage of research, along with concerns that may need to be addressed.

Stage of research: Data analysis

Tool used	Descriptions of use	Opportunities / Concerns
ChatGPT4.0 Plus/ Turbo (code interpreter option)	Python code generation: Using the code interpreter option, you can provide data in a file, then ask for Python code to be generated that will analyze the data. The Python code can be executed. Follow-up prompts can be used to refine the analysis.	The saved Python code can be used to analyse other datasets without uploading any data to ChatGPT. ChatGPT Plus/Turbo is subscription based. Claude is a free alternative. Python code should be checked before reporting. Uploading sensitive data would not be appropriate.
ChatGPT4.0	Advanced data analytics: Can analyse large multimodal models. It generates text outputs (natural language, code, etc.) given inputs consisting of interspersed text and images.	While less capable than an expert in many real-world scenarios, it performs well for many professional and academic tasks. May generate harmful advice, buggy code, or inaccurate information.
Voyant	Text analysis: An open-source, web-based application for performing text analysis. It supports scholarly reading and interpretation of texts or corpora.	It can be used to analyze online texts or ones uploaded by users. Prolonged text-loading time and the challenge of gathering information using some visualization tools.
Maxqda and related data analysis tools	Text thematic analysis: Summarise texts, coding of themes, and identifying common themes across your codings. Can help to explain the coded content.	Can choose the language and the length of the summary. Can modify the summaries as needed.

Publishers associated with universities as well as local journals have also implemented their own policies. This is the [Generative Artificial Intelligence Policy](#) of the South African Orthopaedic Journal. The University of Johannesburg (UJ) Press has introduced an [Artificial Intelligence and Generative AI Policy](#).

In the wider region, the Southern African Regional Universities Association (SARUA) has issued a [statement on ChatGPT and other AI tools](#). Globally, it is worthwhile looking at the OECD's study on the [Emerging governance of generative AI in education](#), as well as the UNESCO [Guidance for generative AI in education and research](#).

11. Guidelines for Using ChatGPT and other Generative AI tools at Harvard

We write today with initial guidelines on the use and procurement of generative artificial intelligence (AI) tools, such as OpenAI's ChatGPT and Google Bard. The University supports responsible experimentation with generative AI tools, but there are important considerations to keep in mind when using these tools, including information security and data privacy, compliance, copyright, and academic integrity.

Generative AI is a rapidly evolving technology, and the University will continue to monitor developments and incorporate feedback from the Harvard community to update our guidelines accordingly.

Initial guidelines for use of generative AI tools:

- **Protect confidential data:** You should not enter data [classified as confidential](#) (Level 2 and above), including non-public research data, into publicly-available generative AI tools, in accordance with the University's [Information Security Policy](#). Information shared with generative AI tools using default settings is not private and could expose proprietary or sensitive information to unauthorized parties.
- **You are responsible for any content that you produce or publish that includes AI-generated material:** AI-generated content can be inaccurate, misleading, or entirely fabricated (sometimes called “hallucinations”), or may contain copyrighted material. Review your AI-generated content before publication.
- **Adhere to current policies on academic integrity:** Review your School's student and faculty handbooks and policies. We expect that Schools will be developing and updating their policies as we better understand the implications of using generative AI tools. In the meantime, faculty should be clear with students they're teaching and advising about their

policies on permitted uses, if any, of generative AI in classes and on academic work. Students are also encouraged to ask their instructors for clarification about these policies as needed.

- **Be alert for AI-enabled phishing:** Generative AI has made it easier for malicious actors to create sophisticated scams at a far greater scale. Continue to [follow security best practices](#) and report suspicious messages to phishing@harvard.edu.
- **Connect with HUIT before procuring generative AI tools:** The University is working to ensure that tools procured on behalf of Harvard have the appropriate privacy and security protections and provide the best use of Harvard funds.
 - If you have procured or are considering procuring generative AI tools or have questions, contact HUIT at ithelp@harvard.edu.
 - Vendor generative AI tools must be [assessed for risk by Harvard's Information Security and Data Privacy office prior to use](#).

It is important to note that these guidelines are not new University policy; rather, they leverage existing University policies. You can find more information about generative AI, including a survey to collect data on its potential use, [on the HUIT website](#), which will be updated as new information becomes available.

12. Generative AI Policy Guidance

Honor Code Implications of Generative AI Tools

The Board on Conduct Affairs (BCA) has been asked to address the Honor Code implications of generative AI tools such as ChatGPT, Bard, DALL-E, and Stable Diffusion. These are novel tools, and both students and instructors have been experimenting with their use in academic settings. While these tools have applications that foster student learning and understanding, these tools can also be used in ways that bypass key learning objectives.

To give sufficient space for instructors to explore uses of generative AI tools in their courses, and to set clear guidelines to students about what uses are and are not consistent with the Stanford Honor Code, the BCA has set forth the following policy guidance regarding generative AI in the context of coursework:

Absent a clear statement from a course instructor, use of or consultation with generative AI shall be treated analogously to assistance from another person. In particular, using generative AI tools to substantially complete an assignment or exam (e.g. by entering exam or assignment questions) is not permitted. Students should acknowledge the use of generative AI (other than incidental use) and default to disclosing such assistance when in doubt.

Individual course instructors are free to set their own policies regulating the use of generative AI tools in their courses, including allowing or disallowing some or all uses of such tools. Course instructors should set such policies in their course syllabi and clearly communicate such policies to students. Students who are unsure of policies regarding generative AI tools are encouraged to ask their instructors for clarification.

The BCA will continue to monitor developments in these tools and their use in academic settings and may update this guidance. Members of the community are encouraged to contact the BCA to provide input, suggestions, and comments on this policy.

NOTE: As part of the BCA's guidance on clear communication of a course's generative AI policy, OCS recommends course instructors provide clear advance notice that they may use detection software to review work submitted for use of generative AI. Other helpful information for faculty and course assistants can be found [HERE](#).

If you are in doubt about whether a generative AI source (or any source) is permitted aid in the context of a particular assignment, talk with the instructor.

13. Getting Started with AI-Enhanced Teaching

Welcome to our guide to leveraging generative AI for teaching at MIT Sloan. The fundamentals of great teaching haven't changed with the emergence of new AI tools. However, if you're struggling to find the time to implement certain research-backed teaching strategies, these new technologies could be just what you need. Here are just a few of the many ways you can use AI in your teaching:

- Do you want to provide students with concrete examples that help illustrate abstract concepts? AI can generate examples on demand.

- Looking to create low-stakes quizzes for comprehension checks? AI can instantly generate practice questions tailored to your needs.
- Want your students to teach new concepts to an inquisitive partner? Consider asking them to converse with an AI model.

This guide will equip you with foundational knowledge, MIT policies, curated tools, ethical considerations, suggested use cases, and avenues to get support when teaching with generative AI tools. Our goal is to provide you with the knowledge and resources to smoothly incorporate these technologies into your teaching.

The Basics

Generative AI is an artificial intelligence subset that learns from data to produce new, unique outputs on a vast scale, ranging from educational content to software code and more. Central to this are foundational AI models trained on massive datasets. Generative AI models are essentially advanced language prediction tools.

There's a lot of jargon involved in discussing generative AI systems. Learn more about generative AI terminology in the [AI Glossary](#).

The following video is the first in [Wharton Interactive's five-part course on Practical AI for Instructors and Students](#). In these videos, MIT Sloan alum and Wharton Associate Professor Ethan Mollick, along with Lilach Mollick, Director of Pedagogy at Wharton Interactive, provide an accessible overview of large language models and their potential for enhancing teaching and learning.

In this first video, you can learn about the following:

- Why AI is now accessible to everyone and how students are using it
- What we mean by AI, specifically large language models and generative AI
- How models like ChatGPT work and their surprising capabilities
- The potentially outsized impact of AI on educators and creative professionals
- Ethical considerations and risks related to generative AI

You can watch the other four videos in the Mollicks' [Practical AI for Instructors and Students Course](#) to learn more about large language models, prompting AI, using AI to enhance your teaching, and how students can use AI to support their learning.

Generative AI Tools

We encourage you to spend some time exploring the generative AI tools in this resource hub. It's important to get a sense of any technology's capabilities and limitations before you integrate it into your teaching. Also, trying these technologies yourself may help you get a sense of how your students are using generative AI.

Before you start using AI tools in your teaching, make sure to review MIT Sloan's [Guiding Principles for the Use of Generative AI in Courses](#).

The tools we've curated in this resource hub fall into these categories:

- **AI Writing and Content Creation Tools:** Large language models accessed through tools like ChatGPT and Claude can help generate written content, provide grammar suggestions, summarize texts, and more. Our overview of AI writing assistants covers the types of support they can provide along with important ethical considerations. While not a substitute for human writing, these tools can help accelerate drafting and revision.
- **AI Data Analysis and Quantitative Tools:** Complex data sets are now more understandable thanks to AI analytics and visualization platforms. Explore options like IBM Watson and ToolsAI to see how algorithms can help process, interpret, and generate insights based on quantitative data. Consider use cases for statistical modeling, data visualization, and other applications while keeping key limitations in mind.
- **AI Image Generation Tools:** Models like DALL-E 3 and Stable Diffusion enable the creation of original images, videos, and other multimedia just by describing desired outputs. With experimentation, they may enable you to transform your visual media workflows.

While you explore each platform's potential, make sure to closely monitor for quality, bias, and responsible usage.

Ethical Considerations

The emergence of powerful generative AI systems presents exciting possibilities for enhancing teaching and learning. However, integrating these technologies into teaching also raises important ethical questions. Three key areas of concern are data privacy, AI-generated falsehoods, and bias in AI systems.

Data Privacy

Make sure to treat unsecured AI systems like public platforms. As a general rule, and in accordance with [MIT's Written Information Security Policy](#), **you should never enter any data or input that is confidential or sensitive into publicly accessible generative AI tools**. This includes (but is not limited to) individual names, physical or email addresses, identification numbers, and specific medical, HR, financial records, as well as proprietary company details and any research or organizational data that are not publicly available. If in doubt, please consult with [MIT Sloan Technology Services Office of Information Security](#).

Note that some of this data is also governed by [FERPA](#) (Family Educational Rights and Privacy Act), the federal law in the United States that mandates the protection of students' educational records (U.S. Department of Education), as well as various international privacy regulations including the [European GDPR](#) and [Chinese PIPL](#).

Microsoft Copilot provides the MIT Sloan community with **data-protected** access to AI tools GPT-4 and DALLE-3. Chat data is not shared with Microsoft or used to train their AI models. Access Microsoft Copilot by logging in with your MIT Kerberos account at <https://copilot.microsoft.com/>. To learn more, see [What is Microsoft Copilot \(AI Chat\)?](#)

Beyond never sharing sensitive data with publicly available AI systems, we recommend that you remove or change any details that can identify you or someone else in any documents or text that you upload or provide as input. If there's something you wouldn't want others to know or see, it's best to keep it out of the AI system altogether (Nield, 2023). This is not just about personal details,

but also proprietary information (including ideas, algorithms or code), unpublished research, or sensitive communications.

It's also essential to recognize that once data is entered into most AI systems, it's challenging—if not impossible—to remove it (Heikkilä, 2023). Always exercise caution and make sure any information you provide aligns with your comfort level and understanding of its potential long-term presence in the AI system, as well as with [MIT's privacy and security requirements](#).

Falsehoods and Bias

There are well-documented issues around AI systems generating content that includes falsehoods (“hallucinations”) and harmful bias (Germain, 2023; Nicoletti & Bass, 2023). Educators have a responsibility to monitor AI output, address problems promptly, and encourage critical thinking about AI’s limitations.

We encourage you to review our resources on protecting privacy, integrating AI responsibly into your course, and mitigating AI’s issues with hallucinations and bias:

- **Navigating Data Privacy:** Using generative AI tools to enhance your teaching requires a strong commitment to data privacy. This article outlines considerations for protecting your and students’ privacy when using publicly available generative AI tools for teaching and learning. These include avoiding sharing sensitive data, treating AI inputs carefully, and customizing privacy settings.
- **Teaching Responsibly with Generative AI:** This guide offers strategies for harnessing AI tools to augment education while addressing AI biases and hallucinations, guiding student engagement with AI tools, and developing AI literacy.
- **When AI Gets It Wrong: Addressing AI Hallucinations and Bias:** This article provides an overview of the biases and inaccuracies currently common in generative AI outputs. It outlines strategies for identifying and mitigating the impact of problematic AI content.

By proactively addressing ethical considerations and AI’s limitations, we can realize the promise of generative AI while upholding principles of fairness, accuracy, and transparency.

AI-Powered Teaching Strategies

Thinking about using generative AI in your teaching but not sure where to start? In this section, we'll walk through several simple strategies for implementing research-based teaching best practices with the help of generative AI tools. These approaches are grounded in the principles of [Universal Design for Learning \(UDL\)](#) and insights from the [learning sciences](#). You can use the strategies as-is or think about creative ways to adapt them to your own courses.

1. Use AI to Generate Concrete Examples

Teaching often involves explaining abstract concepts or theories. While these are essential for academic understanding, they can sometimes be challenging for students to grasp without real-world context. You can use generative AI tools to come up with many concrete examples to make abstract ideas more relatable and understandable for students.

How to Implement This Strategy:

Identify an abstract concept. Select one abstract concept or theory that you'll be covering in your lesson.

Choose a generative AI tool. Select one or several [AI Writing and Content Creation Tools](#) that you'll use for this task.

Teach the AI. Prompt your chosen AI tool to engage with the concept you've selected. If the tool is connected to the internet, you can ask it to look up and summarize the concept. If the tool is not connected to the internet, provide it with open-source content describing the concept and ask it to summarize that information.

Prompt the AI. Ask your chosen chatbot for examples or applications of the chosen concept. You can use a prompt like this one created by Ethan Mollick and Lilach Mollick: "I would like you to act as an example generator for students. When confronted with new and complex concepts, adding many and varied examples helps students better understand those concepts. I would like you to ask what concept I would like examples of, and what level of students I am teaching. You will look up the concept, and then provide me with four different and varied accurate examples of the concept in action" (Mollick & Mollick, 2023-b).

Review and select examples. From the generated examples, select the most relevant and clear examples that align with the lesson's objectives. Always verify the accuracy of the

examples provided by the AI using trusted sources. Make sure to address and eliminate any harmful bias in AI-generated examples.

Integrate the examples into lessons. Incorporate these examples into your lectures, discussions, or assignments.

What's the research? Concrete examples help bridge the gap between abstract theories and real-world applications. Research shows that exploring tangible instances can help students better relate to and understand complex concepts, activating their background knowledge and making learning experiences more meaningful (Smith & Weinstein, n.d.-a; CAST, n.d.-b).

2. Use AI to Create Practice Quizzes

Frequent low-stakes quizzes are a great way to help students test their knowledge and reinforce their understanding. However, creating quizzes can be time-consuming for faculty. With the rise of generative AI tools like ChatGPT, though, it's now possible to streamline the quiz creation process. You can use AI to generate practice quizzes tailored to specific topics. Moreover, these AI-generated quizzes can be adapted to fit various teaching approaches and course requirements, offering a flexible solution for assessment needs.

How to Implement This Strategy:

Choose your topics. Identify the topics or concepts for which you want to create practice quizzes.

Select a generative AI tool. Identify one or several [AI Writing and Content Creation Tools](#) that you'll use for this task.

Teach the AI. Prompt your chosen AI tool to engage with the concept you've selected. If the tool is connected to the internet, you can ask it to look up and summarize the concept. If the tool is not connected to the internet, provide it with open-source content describing the concept and ask it to summarize that information.

Prompt the AI. Ask the AI tool to generate quiz questions related to these topics. Use a quiz-question generating prompt like this one created by Ethan Mollick and Lilach Mollick: "You are a quiz creator of highly diagnostic quizzes. You will look up how to do good low-stakes tests and diagnostics. You will then ask me two questions. (1) First, what, specifically, should the quiz test. (2) Second, for which audience is the quiz. Once you have my answers you will look up the topic and construct several multiple choice questions

to quiz the audience on that topic. The questions should be highly relevant and go beyond just facts. Multiple choice questions should include plausible, competitive alternate responses and should not include an ‘all of the above’ option. At the end of the quiz, you will provide an answer key and explain the right answer” (Mollick & Mollick, 2023-b).

Review and refine the results. Examine the generated questions for relevance and accuracy. Remove any content that perpetuates harmful biases. Modify or refine as necessary.

Distribute the quizzes to students. Share the practice quizzes with students. You may want to incorporate the questions into a [Canvas quiz](#).

What's the research? Retrieval practice, or the act of recalling information from memory, strengthens memory retention (Smith & Weinstein, n.d.-d). Practice quizzes offer students an opportunity to test their understanding and reinforce their learning, making the information more retrievable in the future.

3. Assign Students to Generate Visual Summaries

Visual aids have always been a cornerstone in effective teaching, aiding in comprehension and retention. With the rise of image-generating AI models, we now have new tools in hand to help create these visual aids. In this use case, you'll ask students to craft visual summaries of specific topics, blending both verbal descriptions and AI-generated imagery. This not only deepens their understanding but also fosters creativity and critical thinking as they evaluate and refine the visuals produced by AI tools.

How to Implement This Strategy:

Assign topics. Provide students with specific topics for which they should create visual summaries.

Guide students to explore AI Image Generation Tools. Follow the tips in our article [Teaching Responsibly with AI](#) to set your students up for success with their chosen AI tool. Make sure they are aware of generative AI's [limitations](#) and [privacy implications](#).

Have students create visual summaries. Ask students to find or generate images that they can use to create visual aids for the assigned topics. Encourage students to combine text and visual information to summarize the topic's main points.

Review and discuss students' work. Examine the visual summaries in class, discussing the concepts and clarifying any misconceptions. If this assignment is graded, make sure to grade based on conceptual understanding rather than image quality.

What's the research? Dual coding is when learners interact with content through both verbal and visual information, enhancing memory and understanding (Smith & Weinstein, n.d.-b). This research-backed study strategy aligns with the Universal Design for Learning checkpoint “Illustrate through multiple media” (CAST, n.d.-a). Visual summaries allow students to integrate two forms of information, deepening their comprehension and making the learning experience more engaging.

4. Ask Students to Teach the AI

Deep understanding often comes from the act of explaining. In the realm of education, having students articulate their understanding of a concept can solidify their grasp and highlight areas needing further clarification. With the advent of AI tools like ChatGPT, students now have an interactive platform where they can practice this act of elaboration. By engaging in detailed conversations with the AI, students can receive instant feedback, refine their understanding, and practice the art of explanation.

To see what this strategy can look like in action, check out our blog post: [**Harnessing AI in Finance: Eric So’s Innovative Take on Teaching Value Investing.**](#)

How to Implement This Strategy:

Choose a generative AI tool. Select a free [AI Writing and Content Creation Tool](#) that students can use for this activity.

Introduce your chosen platform to students. Follow the tips in our article [Teaching Responsibly with AI](#) to set your students up for success with their chosen AI tool. Make sure they are aware of generative AI's [limitations](#) and [privacy implications](#).

Assign topics. Provide students with specific topics or concepts they should explain to the AI.

Invite students to interact with the AI. Encourage students to have detailed conversations with the AI, explaining concepts and receiving feedback.

Reflect and discuss. Ask students to reflect on their conversation with the AI and identify areas for improvement.

What's the research? Elaborative interrogation is a research-backed study strategy in which students deepen their understanding by asking questions and explaining concepts (Smith & Weinstein, n.d.-c). By interacting with AI, students can practice this strategy, enhancing their comprehension and reinforcing their learning.

Get Support

As you consider how to best use generative AI in your course, questions will arise. [Contact us](#) today for a personalized consultation. We're here to be your thought partner during your development and implementation process.

Conclusion

Integrating artificial intelligence into your teaching offers both opportunities and challenges. In this guide, we've provided an initial roadmap to begin exploring this new space. We've covered the basics of what generative AI is, considered its potential benefits, and explored practical use cases to incorporate generative AI tools into teaching. We've also emphasized the importance of ethical considerations like prioritizing student privacy and addressing potential biases.

While AI offers powerful tools to augment our teaching methods, the human touch remains irreplaceable. The goal is not to replace educators but to empower them with additional resources. By combining the strengths of AI with the expertise of skilled instructors, we can create richer, more effective learning experiences for our students.

As you move forward, remember that you're not alone on this journey. Our team is here to support you, answer questions, and provide guidance. We're excited to see how you'll harness the potential of AI in your classrooms and look forward to hearing about your experiences. Let's explore, learn, and innovate together.

14. Generative AI Guidance

We encourage faculty to try out or experiment with generative AI (GAI) tools, which can be used to generate ideas, summarize articles, develop computer code, create images, and compose music.

We also encourage faculty to have explicit conversations with students about the permissibility of their use in your courses and their independent work. If you allow students to use GAI tools, we suggest that you set clear and explicit guidelines and help your students understand the risks associated with these programs. In large language models, such as ChatGPT, risks include inaccuracies, fabrications (“hallucinations”), and amplified biases (see more below). We also take seriously the risk that use of these tools may short-circuit learning.

Syllabus Language

Faculty have the discretion to set their own GAI policies (see Jill Dolan’s August 2023 memo). This means that students likely will encounter different rules about AI use in different courses. For this reason, we strongly encourage faculty to articulate a clear policy on AI use in your syllabus. Make clear:

- Whether AI tools are prohibited or permitted in your courses
- Whether AI tools are permitted for certain tasks
- Whether AI tools are permitted for assignments or for certain stages of an assignment
- Which AI tools are permitted in your courses
 - AI tools include Chat GPT, Bard, Grammarly, Github Copilot, Google Translate, Adobe Firefly, etc.
- The use of Generative AI should be acknowledged; see citation guidance from MLA, APA, and Chicago

Keep in mind that a course policy that allows some use of GAI may introduce complexity and open up the possibility of students using GAI in ways you don’t intend.

Below you will find Princeton-specific examples of syllabus language, grouped by category. You may also find this Chronicle article useful as you develop a policy statement. We also highly recommend the guidance and examples our colleagues at Georgetown have curated.

Ethical and Other Risks

We recommend that faculty consider and discuss with students the significant ethical considerations and risks of using generative AI. The most important concerns are:

Equity and Access

Students' varying levels of AI literacy coupled with unequal access to technology and lack of exposure to AI tools exacerbates existing digital divides in education. Safer, more accurate AI tools are often locked behind paywalls, giving rise to concerns regarding affordability and equitable access. Even though today's student population is often well-versed in digital technology, disparities in digital literacy education and skill acquisition can affect students' performance in college.

Student data and privacy

When students create an account in a program, they share personally identifiable information like their email address and phone number. Large language models such as ChatGPT or Bard can store conversations and uploaded content, which they might repurpose as training data. Princeton's Information Security Office has written the following position paper on the Prohibition of University Data in Artificial Intelligence (AI) Solutions.

If you elect to use AI tools that require students to create accounts, we suggest that you highlight these risks and review the data usage policies with your students. Consider, in fact, making this a classroom exercise. You might also offer alternative options for students who are not comfortable creating their own accounts.

Inaccuracies and fabrication

Generative AI fabricates data, invents facts, and produces persuasive but completely inaccurate arguments, according to researchers at Stanford. When used as a research aid, these programs can concoct citations. ChatGPT, for instance, incorrectly stated that Princeton's Hal Foster had written an article called "The Case Against Art History" in October. The citation included volume number, year, and page references—all a fabrication. Making students aware of this tendency toward inaccuracy might help to deter them from relying on these tools.

Cognitive Offloading

Cognitive offloading involves delegating the mental demands of a task to a technology or tool, such as relying on a calculator or smartphone reminders instead of one's own knowledge and abilities. People may offload a task when they think the technology is more capable, they have a high degree of trust in the tools, and the tools are easily accessible. Offloading may improve a student's short-term performance (i.e., getting good grades on an assignment) but diminish their

long-term learning and cognition. We suggest that faculty encourage students to use AI to enhance their learning, not as a replacement for their own cognition.

Bias and stereotypes

Generative AI is fed and trained on data that can be biased and inaccurate, or geographically and racially skewed. It has a tendency to reproduce stereotypes. If prompted to depict a “Native American,” for instance, image-making software like DALL-E 2 and Stable Diffusion tend to produce images of people with traditional headdresses. Or, if asked to illustrate a profession using an adjective like “emotional” or “sensitive,” the program is more likely to produce an image of a woman as this article by the MIT Technology Review demonstrates.

Labor concerns with how AI tools are trained

Companies like OpenAI have relied on labor from the Global South to train their models, requiring workers to read and categorize graphic texts to identify hate speech, violence, and sexual abuse. This source offers a fuller account.

Environmental Impact

The computational requirements associated with large language models like ChatGPT contribute to high rates of energy consumption, carbon emissions, and electronic waste. Researchers at the University of Massachusetts found that training large AI models can produce nearly five times the lifetime emissions of an average car (including fuel). As AI datasets and models grow in complexity, so do their environmental impact.

Implications for Assignments

Generative AI requires us to be very intentional about assignment design -- to maximize students' opportunity to engage critically with course material and minimize their risk of overusing GAI.

Regardless of whether you permit the use of GAI tools, we encourage you to:

- Define your course learning goals, and share them with students.
 - Explain to students what they will learn by completing your assignments. In what ways will it help them develop the skills or master the content of your discipline?
- Include a generative AI policy on your syllabus.

- You might also work with students to set a class policy, as Associate Professor Molly Crockett does in their Psychology courses. See McGraw's Faculty Resource Library for guidance on Creating Community Agreements.
- Test your prompts.
 - To understand more about the strengths and limitations of GAI tools, experiment with your assignment prompts and evaluate the results. For guidance on how to effectively prompt Chat GPT, see Open AI's resource on Prompt Engineering.
- Scaffold assignments.
 - Scaffold students' work with draft and revision deadlines that offer you opportunities to give feedback.
- Incorporate reflection into assignments.
 - Ask students to demonstrate their thought processes and reflect on their work. For example, they might annotate their solution to a problem, write an artist's statement to accompany a submission, or write a cover letter for an essay.
- Assign “creative critical” assignments.
 - Design assignments that ask students to engage creatively as well as critically with course material. This might take multiple forms, including digital assignments like digital exhibitions, podcasts, or story maps. Even without the use of digital technology, consider assignments that ask students to riff on, mix up, or playfully and purposefully engage course material. We have many ideas to share with you; feel free to consult with us.
- Try oral assignments, especially if you do not permit the use of GAI.
 - Devise oral assignments such as presentations, simulations, or role plays. These can be low-stakes activities—for example, asking a student to talk through their response to a problem or share ideas as part of a “fishbowl” discussion—or higher-stakes activities that require advanced planning and preparation.
- Make an appointment for a consultation with us.
 - We're very happy to help you think through how generative AI may affect your teaching. We offer consultations in person and over Zoom; be in touch with us at mcgraw@princeton.edu.

Assigning Generative AI

If navigating AI is a skill you think is important for students to develop, you might design activities and assignments that embrace it. If you do ask students to use GAI tools, be mindful of the ethical concerns and other risks that they present. Remember that some students may have access to subscription-based tools like Chat GPTPlus, while others will only have access to the less powerful free versions. Be prepared to offer alternative assignments or other workarounds to students who don't feel comfortable using these tools—which often require students to create an account—themselves.

- Ask students to analyze its output. For example, after they complete their own drafts of an assigned essay, you might ask students to request a draft of the assignment from a generative AI tool and analyze and/or critique the work it produces. Jacob Shapiro, Professor of Politics and International Affairs, requires students to prompt Chat GPT and then share the responses with classmates to revise them. Associate Professor Alexander Glaser from Mechanical and Aerospace Engineering asks students to compare their answers to those composed by ChatGPT and reflect on the differences between responses from a human and those from a machine. Steven Strauss, Visiting Professor in SPIA, asks graduate students to grade ChatGPT's response to a prompt and then reverses the process, asking students to submit their draft answers to ChatGPT (with the appropriate context) so it can give them feedback.
- Allow students to use the tool for one part of a larger assignment. For example, Heather Thieringer, University Lecturer in Molecular Biology, allows students to use ChatGPT to create a potential introduction to their lab report, which they include as an appendix. The students critique and correct the response as part of the assignment.
- Emphasize the skill of prompt engineering. Assign students to use the tool and to turn in the prompts they use to get their responses. Ask them to write a short paper reflecting on how altering their prompt changed the output.
- Use the tools to enhance students' creativity. In his *Storytelling* course, Professor of Slavic Languages and Literatures Yuri Leving asks students to illustrate their writing projects with images produced by an AI generator. He also asks students to write stories inspired by images he has generated using the tool, giving them experience both creating images from text and generating text from images.

- Ask students to analyze the benefits and drawbacks of generative AI for certain tasks in class discussions, debates, or written assignments. For example, Steven Strauss, Visiting Professor in SPIA, devotes class time to what he calls “GAI housekeeping” before requiring students to use the tools, addressing topics such as student accountability, algorithmic bias, the potential for hallucinations, ethical dilemmas, and replicability concerns. Once students understand the challenges and limitations inherent in GAI technology, Strauss asks them to make and support an argument about how ChatGPT and similar tools might be used to improve productivity on an everyday task.

Faculty have expressed interest in hearing about how colleagues are using AI tools in their courses. If you are assigning GAI in your course and would be willing to share your assignment, please reach out to us at mcgraw@princeton.edu.

Detection Software and Academic Integrity

Though companies like Turnitin, ZeroGPT, and OpenAI have all developed AI detection capabilities, we do not recommend you use such software to attempt to determine if student work is AI-generated. Our recommendation against using these tools is based both on Princeton’s standards for academic integrity and the practical limits of these tools. Detection tools seem unreliable at best and biased at worst. The creators of these tools have warned against using them to make decisions about academic honesty. Research has also demonstrated that the software consistently misclassifies writing samples by non-native English writing as AI-generated.

Instead, we encourage you to emphasize your learning goals, consider our guidance on assignment design, and include a clearly stated GAI policy on your syllabus.

If you suspect a student has used an unauthorized GAI tool in an assignment, please contact Joyce Chen at jgchen@princeton.edu or 609-258-3054. If you suspect a student has used an unauthorized GAI tool in an exam, please contact the Honor Committee at honor@princeton.edu.

Teaching guidance and case reports on teaching with GAI

1. Assigning AI: Seven Approaches for Students, with Prompts (**Social Science Research Network**)
2. ChatGPT and the Rise of Generative AI: Threat to Academic Integrity? (**Science Direct**)

- 3. Artificial Intelligence and the Future of Teaching and Learning ([U.S. Department of Education](#))**
- 4. Artificial Intelligence and Education: A Reading List ([JSTOR Daily](#))**
- 5. Reactions: Princeton faculty discuss ChatGPT in the classroom ([The Princetonian](#))**
- 6. The AI Pedagogy Project ([Harvard University](#))**
- 7. Learn With AI Toolkit ([University of Maine](#))**
- 8. Teaching in the Age of AI ([Vanderbilt University](#))**
- 9. Artificial Intelligence Teaching Guide ([Stanford University](#))**
- 10. Intentional Pedagogy with AI Technology ([Brown University](#))**
- 11. Why I'm Encouraging my Students to use Generative AI (e.g., ChatGPT) When Writing Their Assignments ([Medium](#))**
- 12. Embracing Generative AI (GAI) in Education: Some Personal Reflections ([Medium](#))**
- 13. Teaching CS50 with AI ([SIGCSE](#))**
- 14. Developing AI Standards of Conduct as a Class ([Exploring AI Pedagogy](#))**
- 15. What I Learned From an Experiment to Apply Generative AI to My Data Course ([EdSurge](#))**

Regularly updated sources on GAI: channels; podcasts; listservs; substacks

- 1. Arvind Narayanan**
- 2. Dr Philippa Hardman**
- 3. One Useful Thing**
- 4. Exploring AI Pedagogy**
- 5. AI+EDU=Simplified**
- 6. AutomatED: Teaching Better With Tech**
- 7. AI in Education**

Links to PU faculty research and initiatives on GAI

- 1. Princeton Language and Intelligence**
- 2. Future Values Initiative**
- 3. Princeton Center for Information Technology Policy**
- 4. Princeton Precision Health**
- 5. Princeton Dialogues on AI and Ethics**

University Report and Resources

1. Generative AI Working Group Report (Office of the Dean of the College)
2. Generative AI (Princeton University Library)

15. Guidance for Syllabus Statements on the Use of AI Tools

Guidance for Syllabus Statements on the Use of AI Tools As Autumn Quarter approaches, the University teaching community is thinking through how to address the use of AI tools, such as ChatGPT, in their courses. Not unlike the disruptions posed by the COVID-19 pandemic, one of the essential strategies for responding to this development is clear, transparent communication with students about our expectations. With the rapid onset of availability and sophistication of AI tools over the past year, students have only had a short amount of time to discover the utility and limits of these tools for their learning as well as expectations of their use in their courses of study. Moreover, the expectations around AI tools will likely vary from one course to the next, making clear and consistent communication of expectations for your course particularly important. This document provides considerations that might shape how instructors craft a syllabus statement about AI tools, followed by example statements covering a range of approaches. The staff in the Chicago Center for Teaching and Learning (CCTL) are always available for one-on-one consultations on using AI tools in your courses, communicating to students about their use of these tools, and writing syllabus statements. To schedule a consultation please email: teaching@uchicago.edu.

Considerations and Examples of Syllabus Statements on the Use of AI Tools

Below are some examples of syllabus statements prepared by the CCTL to help guide your writing of your own statement on the use of AI tools in your courses. This is not an exhaustive list and the statements are generic in nature, with the idea that these examples will provide some ideas and starting points for thinking about how you might want to communicate your own approach on the role of AI tools in your courses. Please feel free to adapt and modify these statements to your particular context and needs. As you draft your own syllabus statement, here are a few considerations to keep in mind:

- Communicate clearly and specifically when AI tools are and are

not allowed, and what uses constitute a violation of academic integrity. • When AI tools are permitted, communicate when and how they should be correctly attributed. o APA guide for citing ChatGPT o Chicago Manual of Style citation guidelines for AI tools o MLA guide for citing AI tools • As with any course policy, providing reasoning that connects it to supporting the learning process helps students to understand the pedagogical rationale for the policy. Communicate to students how the use of AI tools do or do not support the learning goals for the course. In many courses, the development of foundational skills and knowledge that are integral to a student's academic and personal growth will need to be developed without the use of AI tools. In other cases, the development of skills and knowledge around the use of AI tools in particular fields may be an important learning goal. • Consider linking to or referencing the student policy on Academic Honesty and Plagiarism. Every new student at the University is provided with the University's Student Manual of University Policies & Regulations. Under the policy, AI tools that are not explicitly allowed by the instructor(s) of a course, will be considered a violation of academic integrity. While it is the responsibility of each student to make themselves familiar with these policies, it can be helpful to remind them of the specific policy on Academic Honesty and Plagiarism. • Consider how word choice and the tone of your statement will communicate and foster trust in your students as they navigate the availability and use of these new technologies. In general, most instructors will allow or limit the use of AI tools in ways that fall under four broad categories: (1) use prohibited in all situations; (2) use with prior permission; (3) use with proper citation; and (4) free use with no citation required. Below are examples of general statements in each of these categories. These examples are best used as starting points to adapt to your teaching, assignments, course design, and style of communicating with your students.

(1) Use prohibited in all situations:

"In this course, we will be developing skills and knowledge that are important to discover and practice on your own. Because use of AI tools inhibits development of these skills and knowledge, students are not allowed to use any AI tools, such as ChatGPT or DallE 2, in this course. Students are expected to present work that is their own without assistance from others, including automated tools. If you are unclear if something is an AI tool, please check with your instructor. Using AI tools for any purposes in this course will violate the University's academic integrity policy. I treat potential academic integrity violations by [...]"

(2) Use with prior permission:

“Students are only allowed to use AI tools, such as ChatGPT or Dall-E 2, on assignments in this course when advance permission is given by the instructor. Students must submit a written request with an explanation of how they will use a particular tool in their assignment. Students are not permitted to use these tools until permission is granted in writing. The instructor may encourage and give permission to students to use AI tools during class activities and in other contexts when it is considered in support of the course learning goals. Unless given permission to use those tools, each student is expected to complete each course assignment without substantive assistance from others, including AI tools. If you are unclear if something is an AI tool, please check with your instructor. Unauthorized use of AI tools for any purposes in this course will violate the University’s academic integrity policy. I treat potential academic integrity violations by [...]”

(3) Use only with proper citation:

“The use of AI tools, such as ChatGPT or Dall-E 2, for this course is allowed for specific assignments only when determined to be in support of the course learning goals. Assignments in which AI tools are permitted will be clearly identified by the instructor and noted in the assignment directions. You are not required to use AI tools, but if you choose to use them for any part of the assignment (from brainstorming to text editing), you must use proper citation (please use APA citation format). Failure to properly cite AI tools is considered a violation of the University of Chicago’s Academic Honesty and Plagiarism policy. If you are unclear if something is an AI Tool, please check with your instructor. I treat potential academic integrity violations by [...]”

(4) Free use with no citation required:

“In this course, students are allowed to use AI tools (such as ChatGPT) on all assignments. No citation is required.”

Additional Resources

To schedule a one-on-one consultation with a member of the CCTL staff, please email: teaching@uchicago.edu If you are interested in reading more, below are some curated resources to learn more about AI tools and how others in higher education are approaching their use in their courses:

- “Should You Add an AI Policy to Your Syllabus? What to consider in drafting your own course policy on students’ use of tools like ChatGPT.” by Kevin Gannon
- “Sentient Syllabus Project”
- “My Assessments Next Semester – Modified for Avoiding & Embracing AI” by Maha Bali
- Crowdsource Classroom Policies for AI Generative Tools

16. Considerations for AI Tools in the Classroom

Considerations for AI Tools in the Classroom

Given the rapid pace of technological innovation and development, higher education, like nearly all industries, is continuously called upon to consider creative approaches to teaching and learning. The following resource offers instructors a brief introduction to Artificial Intelligence (AI) Tools, specifically ChatGPT, along with several strategies they might consider for navigating or engaging with these tools in their courses.

For instructors: share your thoughts with us!

Have additional questions about AI Tools in your classroom? Have examples of how you are integrating these tools in your course or redesigning your assignments to address them? Email CTLFaculty@columbia.edu to share your questions and innovations!

Cite this resource: Columbia Center for Teaching and Learning (2023). Considerations for AI Tools in the Classroom. Columbia University. Retrieved [today's date] from <https://ctl.columbia.edu/resources-and-technology/resources/ai-tools/>

What are AI Tools?

While instructors will continue to encounter new tools and technological innovations, which can sometimes feel overwhelming, the science of teaching and learning offers support for navigating and responding to innovative tools, expanded opportunities, and ongoing shifts.

Most recently, higher education has grappled with the seemingly-overnight introduction of ChatGPT, which first became available in November 2022 and has since captivated the discourse in higher education with its ongoing evolution and iterations. The following section offers an overview of ChatGPT, what it does, and its current limitations. It's important to recognize, however, that because the technology is self-learning, it can improve and evolve quickly; thus, what is a limitation today could be addressed in future iterations or versions. This was perhaps

made most evident with the release of GPT-4 during the Spring 2023 semester, which introduced a series of new capabilities for the generative AI tool. Furthermore, since the release of ChatGPT hundreds of AI apps have exploded into the scene, some of which assist in research (e.g. consensus.app and scite.ai), brainstorming (e.g. mymap.ai), and student presentations (e.g. lenovo.ai, twelvelabs.io, and invideo.io).

ChatGPT

ChatGPT, created by OpenAI, is a “model [...] which interacts in a conversational way” and has the ability to “answer followup [sic] questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests” ([OpenAI Blog, “ChatGPT”](#)). In respect to the classroom, ChatGPT can produce written responses to input prompts, write essays and poems, assist with computer code, provide feedback on student-written text, and more.

The most recent version, GPT-4, is described as “more creative and collaborative” with the ability to “generate, edit, and iterate with users on creative and technical writing tasks” ([OpenAI, GPT-4](#)). Unlike its predecessors, GPT-4 is capable of producing longer form responses, with more complex analyses in response to particular prompts. It has wider capabilities such as image analysis, and the addition of plug-ins that connect ChatGPT to other third party services, including information from the web. It is important to note, however, that GPT-4 and its capabilities are only available through a *paid subscription* to GPT Plus; these more advanced features are not built into the free uses of ChatGPT.

Despite these expanded capabilities, like many AI tools, ChatGPT does have a variety of limitations. Some of its current limitations include:

- Sharing incorrect information: The paid GPT-4 version of ChatGPT has access to the internet, while GPT-3.5 (the free version) does not and can only access information prior to 2021. Since ChatGPT is a predictive text model, it can at times make up information, especially if the information is not easily accessible. This makes ChatGPT prone to making factual mistakes, often confusing similar information, and even making up citations when asked to produce them.

- Generating personal reflections: ChatGPT can generate responses to questions and prompts, but it is still an AI bot; thus, it is unable to respond to a prompt that requires a student's personal experience or reflections.
- Producing non-text based responses: Despite the availability of other image-based AI generators such as Dall-E 2 and MidJourney, ChatGPT responses are strictly text-based. Therefore, any response produced can only offer text. With its most recent updates, ChatGPT can, however, accept image input prompts to generate captions and produce image analyses.

An awareness of ChatGPT's capabilities and limitations can help instructors talk openly with their students about the potential role of the tool in the course and what is expected in terms of student engagement with AI tools. Additionally, knowing what ChatGPT can and cannot do can also help instructors make decisions around course and assignment design. The following section offers several approaches for instructors to consider when developing their own approaches to ChatGPT in their own courses. Instructors can explore OpenAI's documentation [Educator considerations for ChatGPT](#) as they consider their own approach to teaching and learning with ChatGPT.

ChatGPT in the Classroom

There is no one correct way to navigate the use of ChatGPT in the classroom. The role it plays (or doesn't) in a course will depend on several factors including course objectives and goals, disciplinary skills that students need to practice, and instructors' own comfort level with engaging with the tool. No matter how instructors respond to the evolution of AI tools, it's important to have explicit expectations, and provide opportunities for clear conversations with students about those expectations. Read on to learn more about several possible approaches and considerations to help instructors start navigating, leveraging, and possibly even engaging ChatGPT (and future AI tools) in their course.

View recording of the “Teaching and ChatGPT” forum held on February 13, 2023. Columbia contributors shared resources and facilitated an informal conversation about AI tools, specifically ChatGPT, and its implications in the classroom.

Develop Course Policies that Include Digital Transparency

It is important to be explicit with students about the expectations around the usage of ChatGPT and other AI tools in your course. For example, ChatGPT is capable of reading a student's essay and providing meaningful feedback that can then be used by the student to make edits. As the instructor, be clear about these expectations: can students use the tool for feedback on their own writing? If so, how should they disclose their use? As with all course policies, especially those around [academic integrity](#), it is essential for instructors to be explicit and transparent with their expectations, and to have frank conversations with their students. Some colleagues are collecting and maintaining [an open-source repository](#) of sample digital transparency language and policies from higher education. While not affiliated with Columbia, this collection can offer some insight into the different approaches institutions and individual instructors are taking when it comes to addressing AI tools in the classroom.

Since ChatGPT's introduction, there has been a parallel rise in tools claiming accurate detection of AI-generated work. As with any form of detection software, there are risks of misidentification, which can have consequences in the classroom. These products are best used with careful consideration, and as one of many ways to work with students. It is also important to include the use of these tools in any discussion with students around course policies, making clear why and how such services may be used in the course.

When developing usage policies in a course, instructors might also consider [partnering with their students](#) to develop the policies. This partnership can create opportunities for instructors and students to talk in detail about the evolution of particular tools, their potential benefits in specific disciplines, and their limitations. Instructors should be explicit about the course objectives and how the use of these tools might interfere with students' learning and their achievement of particular learning goals.

Scaffold Activities and Assignments

Regardless of AI tool innovation or evolution, one important approach for instructors is to leverage scaffolded activities and assignments. Scaffolding is the process of breaking down a larger

assignment into subtasks, which create opportunities for students to check-in and receive feedback. At the same time, scaffolding can help instructors become more familiar with students' work as the semester progresses. This cyclical process of feedback and revision makes the use of tools like ChatGPT challenging and perhaps even unlikely, as students will provide drafts incrementally, and engage in the process of drafting and revision. More importantly, though, this breaking down of a large project into incremental parts helps students to more deeply engage with the different skills and component parts, while also creating valuable time for feedback and reflection throughout the process. For ideas on how to scaffold student work, view the [recording](#) of "Using AI writing tools in your scientific writing process" (May 2, 2023) in which Tim Requarth shares how he uses AI tools to streamline and support the writing process.

Design Authentic Assessments for Learning

Authentic assessments centered on student learning can help instructors and learners make intentional choices about the integration of AI tools into the writing process. These kinds of authentic assessments ask students to apply the course concepts they have learned to a "real world" situation or problem. In doing so, authentic assessments can enhance student learning by engaging students in "doing" a particular subject and practicing specific disciplinary skills that will help prepare them for their professional lives outside the classroom. Additionally, these kinds of assignments engage students in higher order thinking and require that they grapple with real world problems and challenges. Designing authentic assessments can ask students to draw from and engage with specific course materials, explore their local community, make connections between course concepts, or also ask students to incorporate their personal experiences or reflections. In this way, authentic assessments can help prevent the use of tools like ChatGPT in that their very design and objectives are rooted within specific course concepts, while also asking students to infuse their own experiences and reflections.

Incorporate AI Tools into Assignment Design

For some courses, depending on the goals and objectives, instructors might consider ways to incorporate AI tools in their assignment design; in doing so, instructors can provide students with opportunities to practice and foster the digital literacy skills they will need for the future. These kinds of [creative assignments](#) might ask students to produce AI-written texts as a way to develop awareness of voice, authorship, and accuracy. Additionally, students could apply a rubric and offer feedback on AI-produced texts to build deeper awareness of a course prompt. Lastly, in some instances, ChatGPT might be called upon as a learning support tool, where students ask for feedback on their own texts, have readings summarized, create personalized study materials, or brainstorm for ideas. No matter the assignment design approach taken, instructors should offer opportunities to discuss the assignment with students, asking them to [reflect](#) on the experience and analyze their engagement with the tool.

Conclusion

Perhaps unsurprisingly, higher education, like all industries, will continue to feel the impacts of technological evolution and growth. As such, and like they always have, classrooms will continue to remain flexible and responsive to this evolution. With digital innovation and developments, the capabilities (and limitations) of today's AI tools, including ChatGPT, will shift and evolve. For that reason, trying to completely ignore or shut out these tools, or even adopting an approach of complete disengagement, will not serve instructors and their students in the long-term. Instead, instructors have an opportunity to rethink and focus on the elements of their course over which they have the most control, including transparent course policies, explicit communication, partnerships with students, and course and assignment design. Leveraging these aspects of teaching and learning can better serve instructors and their students no matter the digital innovations of the future.

The CTL is here to help!

Looking to develop scaffolded authentic assignments for your course? Want to know more about designing assignments that incorporate AI Tools? Have questions about digital transparency in

your course policies? The CTL is here to help – email CTLFaculty@columbia.edu to schedule a 1-1 consultation!

Columbia CTL Related Resources

- [Designing Assignments for Learning](#)
- [Feedback for Learning](#)
- [Getting Started with Creative Assignments](#)
- [Learner Perspectives on AI Tools: Digital Literacy, Academic Integrity, and Student Engagement](#)
- [Metacognition](#)

Additional Related Resources

- Barnard Center for Engaged Pedagogy (2023). [Generative AI & the College Classroom](#).
- Cornell University Center for Teaching Innovation (2023). [Promoting Academic Integrity in Your Course](#).
- Digital Futures Institute, Teachers College, Columbia University (2023). [ChatGPT and Other Artificial Intelligence \(AI\) in the Classroom](#).
- Digital Futures Institute, Teachers College, Columbia University (2023). [Thinking About Assessment in the Time of Generative Artificial Intelligence](#).
- Montclair State Office for Faculty Excellence. (2023). [Practical Responses to ChatGPT](#).
- Penn Center for Teaching and Learning (2023). [ChatGPT and Its Implications for Your Teaching](#).
- University of Central Florida Faculty Center. (2023). [Artificial Intelligence Writing](#).
- Yale Poorvu Center for Teaching and Learning (2023). [AI Guidance](#).

17. Guidance on the Use of Generative AI and Large Language Model Tools

Guidance on the Use of Generative AI and Large Language Model Tools

Caltech provides this initial guidance to encourage the responsible use of generative artificial intelligence (GenAI) and large language model (LLM) tools and technologies, such as OpenAI's ChatGPT and Dall-E and Google's Bard in research, education, and administrative work at Caltech. As a research and education institute, committed to advancing the frontiers of science and engineering and expanding knowledge, we support a responsible, measured experimentation with and use of new technologies. While doing this, however, Caltech requires that you follow all existing applicable regulations and Institute policies. These include, but are not limited to, ensuring protection of confidential, personal, or business information and intellectual property, and adherence to the honor code, course requirements, research integrity, and publication ethics.

GenAI and LLM technologies have evolved rapidly in their use and application this past year and are expected to continue to evolve in ways society cannot predict. Likewise, our guidance for the appropriate use of these tools is written for the present moment and will likely evolve alongside the technology. In the interim, however, as you use GenAI and LLM technologies in your work at Caltech, we ask that you apply these four guiding principles to your practice: disclosure, data and information protection, content responsibility, and Caltech's honor code.

1. Disclosure: When using GenAI, always disclose promptly, or reference the use of GenAI tools and application plug-ins, as applicable. This transparent disclosure ensures that others are aware when GenAI was used to generate content and reduces misunderstandings regarding the source of information, potentially limiting claims of academic dishonesty or plagiarism. When using GenAI to write or publish, please make sure to follow the guidance provided by the course instructor or journal or manuscript publisher/editor. For example, some may require that the GenAI be included as an author, others may simply require acknowledgement.
2. Data and Information Protection: Federal, state, and local laws as well as Caltech policies may limit data that can be disclosed. Unless you are using a GenAI application that ensures separation of your entry from other entries and confidentiality (usually a paid service), uploading content into GenAI (Open GenAI), is a public disclosure. It is safest to assume data or queries uploaded into Open GenAI tools will become public information, unless otherwise indicated.

In order to protect Caltech data and information, do not enter, contribute, or otherwise input sensitive, confidential, or restricted information into open GenAI tools. This includes, but is

not limited to, data covered by regulations such as FERPA and HIPAA, any intellectual property or unpublished research data, export-controlled data, and other sensitive HR, business, or administrative data. Caltech is considering a subscription to a restricted GenAI and will keep you apprised of its progress in securing such a service. In the meantime, Caltech has and continues to reserve the right to disable or limit access to AI companion tools in enterprise business software and applications, such as Zoom and Microsoft Office suites.

3. Content Responsibility: Remember that GenAI systems are fallible. Responses can be inaccurate, misleading, and even entirely fabricated. Therefore, you should always review and assess all output generated by GenAI tools for accuracy before relying on them or distributing them publicly.
4. Honor Code: Caltech's honor code underscores the importance of ethical conduct and fairness and extends to the use of GenAI tools and is stated as follows: "No member of the Caltech community shall take unfair advantage of any other member of the Caltech community."

Please note that we offer these guidelines in addition to the teaching resources that have already been provided by the [Center for Teaching and Learning and Outreach](#).

The Institute's guidance promotes responsible and ethical use of GenAI and LLM tools at Caltech, and fosters a community that values transparency, integrity, privacy, accuracy, and fairness. Caltech may update its guidance as the technology and regulatory and commercial landscapes evolve. Thank you for adhering to these guidelines. If you have any questions about the use of this dynamic technology, please email gen_ai@caltech.edu.

18. Appropriate use of ChatGPT and Similar AI Tools

With the emergence of ChatGPT and other AI tools, many members of our community are eager to explore their use in the university context. This advisory provides guidance on how to use these tools safely, without putting institutional, personal, or proprietary information at risk. Additional guidance may be forthcoming as circumstances evolve.

Allowable Use:

- Publicly-available information ([Protection Level P1](#))
- ([link is external](#))
-) can be used freely in ChatGPT.

- In all cases, use should be consistent with [UC Berkeley's Principles of Community](#)
- (link is external)

Prohibited Use:

- At present, any use of ChatGPT should be with the assumption that no personal, confidential, proprietary, or otherwise sensitive information may be used with it. In general, [student records subject to FERPA](#)
- (link is external)
- , and any other information classified as [Protection Level P2, P3, or P4](#)
- (link is external)
- should not be used.
- Similarly, ChatGPT should not be used to generate output that would be considered non-public. Examples include, but are not limited to, proprietary or unpublished research; legal analysis or advice; recruitment, personnel or disciplinary decision making; completion of academic work in a manner not allowed by the instructor; creation of non-public instructional materials; and grading.
- Please also note that OpenAI explicitly forbids the use of ChatGPT and their other products for certain categories of activity, including fraud and illegal activities. This list of items can be found in their [usage policy document](#)
- (link is external)

Additional Guidance:

For further guidance on the use of ChatGPT for teaching and learning, please see [Understanding AI Writing Tools and their Uses for Teaching and Learning at UC Berkeley](#)
(link is external) from Research, Teaching & Learning.

Rationale for the Above Guidance:

UPDATE 7/1/2023: The University of California recently renegotiated the UC systemwide agreement with Microsoft to include Microsoft Azure OpenAI. UC Berkeley is currently working with Microsoft to establish how to use this service under UC's agreement. [Note: For questions regarding the approved use of Microsoft's Azure's Open AI service, please consult with the Privacy Office at privacyoffice@berkeley.edu

(link sends e-mail)

As of September 2023, the University of California's agreements with the parent companies of ChatGPT and other generative AI services, which include our terms and conditions, Appendix Data Security and other privacy protections, do not cover the use of ChatGPT or other similar generative AI services. The UC Office of the President is working on this issue. We hope to see this addressed in the near future and will update this guidance when additional information is available.

Personal Liability: ChatGPT uses a click-through agreement. Click-through agreements, including OpenAI and ChatGPT terms of use, are contracts. Individuals who accept click-through agreements without delegated signature authority may face personal consequences, including responsibility for compliance with terms and conditions. [1]

Guidance on Appropriate Use

For questions regarding the appropriate use of ChatGPT and other AI tools, please contact privacyoffice@berkeley.edu

(link sends e-mail)

References

- [Educator considerations for ChatGPT](#)
● (link is external)
- [OpenAI sharing & publication policy](#)
● (link is external)
- [OpenAI usage policies](#)
● (link is external)
- [OpenAI privacy policy](#)

- (link is external)
- [OpenAI terms & policies](#)
- (link is external)
- [1] Delegations of Authority: To find out who has signature authority at Berkeley to “sign” a click-through agreement, go to the Delegations of Authority webpage in the Office of the Chancellor Compliance Services website.

Additional Readings

- [Governor Newsom Signs Executive Order to Prepare California for the Progress of Artificial Intelligence](#)
- (link is external)
-
- [Understanding AI Writing Tools and their Uses for Teaching and Learning at UC Berkeley](#)
- (link is external)
- , from UC Berkeley Research, Teaching & Learning
- [Quantamagazine: The Unpredictable Abilities Emerging From Large AI Models](#)
- (link is external)
- [The Atlantic: Don’t Be Misled by GPT-4’s Gift of Gab](#)
- (link is external)
- [University of California Presidential Working Group on Artificial Intelligence](#)
- (link is external)
- [Inclusive Intelligence: Artificial Intelligence in the Service of Science, Work, and the Public Good](#)
- (link is external)
- , from UC Berkeley Research

19. Guidelines for the Use of Generative AI Tools

Dear Members of the Yale Community,

Publicly available generative Artificial Intelligence (AI) tools such as ChatGPT, Bing, Midjourney, and Bard have garnered tremendous attention in the past year. The field of generative AI is developing rapidly. While its precise impacts are unknown, this technology will transform how we learn, teach, conduct research, and carry out daily tasks. We encourage you to experiment with AI tools. As you explore AI's potential, please adhere to the following general guidelines, which align with existing university policies and uphold our institutional commitment to safety, security, and academic integrity.

1. **Protect Yale's confidential information and your own.** Do not enter confidential or legally restricted data or any data that Yale's data classification policy identifies as moderate or high-risk into an AI tool. If you are not sure whether you should share certain data, please review [Yale's data classification policy](#).
2. **Assume all information shared will be made public.** Treat all information shared with an AI tool as if it will become public. Do not share information that is personal or sensitive, and be mindful that the information you input into an AI tool may be retained.
3. **Always follow academic integrity guidelines and institutional standards of conduct.** All students and faculty are expected to know and adhere to their school's academic integrity policies. Faculty members are expected to provide clear instructions on the permitted use of generative AI tools for academic work and requirements for attribution. Likewise, students are expected to follow their instructors' guidelines about permitted use of AI for coursework.
4. **Be alert for bias and inaccuracy.** AI-generated responses can be biased, inaccurate, inappropriate, or may contain unauthorized copyrighted information. We are each responsible for the content of our work product. Always review and verify outputs generated by AI tools, especially before publication.
5. **Protect yourself and your credentials.** Never share your Yale NetID and password with AI tools, and always be aware of phishing schemes. For information, tips, and toolkits on cybersafe practices, visit [Yale's Cybersecurity website](#), which also includes information about security

policies and standards.

6. **Seek support.** The university is working to support procurement practices that coordinate shared interests and minimize institutional risk. If you are considering acquiring an AI product, please conduct an initial review of the tool to ensure that it conforms to institutional security requirements. Use [Yale's purchasing intake portal](#) or contact purchasing.helpdesk@yale.edu.

Generative AI is evolving quickly, as is institutional support for utilizing these tools. For up-to-date recommendations on teaching and learning, please visit the [Poorvu Center's webpage on “AI Guidance.”](#) Detailed AI guidelines for staff are available on Yale’s [data governance website](#).

As a university community dedicated to exchanging ideas, disseminating knowledge, and fostering a climate for breakthrough discoveries, we will embrace technological tools and harness their power for innovation. We must do this safely and responsibly. We thank you for your efforts in adhering to these guidelines.

20. Statement on Guidance for the University of Pennsylvania Community

Penn embraces innovations like generative artificial intelligence (“AI”) models in teaching, learning, research, and the effective stewardship of Penn’s resources. To this end, this document provides guidelines for members of the Penn community who are using, or interested in using, AI in pursuit of Penn’s mission.

Scope

This document is scoped to generative AI using large language models provided by third parties. Generative AI describes algorithms, such as ChatGPT and other large language models, that can be used to create new content, including text, code, and simulations.

This statement is not intended as legal advice or an exhaustive set of best practices and should not be viewed as a final policy. The AI field is rapidly evolving in terms of technology, deployment models, third-party relationships, terms of service, regulatory landscape, and academic-industry

partnership structures. It is anticipated that this document will be updated regularly and interact with other sources of policy, ethics, and governing legal authority.

General Guidance for Penn Community (Educators, Staff, Researchers, and Students)

Transparency

Be transparent about the use of AI. Disclose when a work product was created wholly or partially using an AI tool and, if appropriate, how AI was used to create the work product.

Accountability

The user of AI should endeavor to validate the accuracy of created content with trusted first party sources and monitor the reliability of that content. Users are accountable for their use of content created by AI and should be wary of misinformation or “hallucinations” by AI tools (e.g., citations to publications or source materials that do not exist or references that otherwise distort the truth).

Bias

When using AI, keep in mind that these tools are often trained on large, unmoderated bodies of text, such as text posted to the internet. This can result in the production of biased and other unintended content. The ability to avoid such biased content is still in the early stages of development.

Privacy & Contracts

Most AI tools and services use input and data from users of the tool to train the model. Additionally, existing tools may incorporate AI features in their service offerings. For this reason, users of AI should avoid sharing personal or sensitive data with the tool and should not input moderate or high-risk Penn data as defined by the [Penn Data Risk Classification](#), or intellectual property, without:

1. Careful consideration and understanding of the tool’s use of Penn data and the service provider’s stated rights to the data, including, but not limited to whether the service provider offers the option to opt-out of using customer’s data to train the AI;
2. A contract in place to protect Penn data; and
3. Review by Penn’s Privacy Office and consultation with the Office of Information Security as coordinated by Procurement when moderate or high-risk data is involved.

Consultation with the Penn Center for Innovation, where intellectual property is involved.

Patient Privacy Protection

It is not permissible under the Health Information Portability & Accountability Act (HIPAA) or Penn Medicine policy to share patient or research participant information in connection with open or public AI tools and services, such as ChatGPT. This is because, as currently configured, such open or public tools and services can use and share any data without regard to HIPAA restrictions and other protections. Therefore, individual patient data and patient data sets (even if deidentified) may not be exposed to open or public AI tools or services, absent institutional approval.

Security

When using AI to write computer code or when creating new technology that leverages AI it is important to be aware of the new kinds of cyberattacks that are being used against AI users. Review the [Office of Information Security guidance](#) on these risks or consult with the Office of Information Security if in doubt.

Data Scraping

The rise of AI models has led to a significant increase in individuals and organizations scraping (i.e., copying) information posted on the internet for the purpose of training new AI models. Be aware that any data posted publicly will likely be scraped and used in this way by third parties. Similarly, while these practices are common, their legality and the potential consequences of these actions are currently being developed but remain unresolved at the time this guidance was issued.

Intellectual Property

Members of the Penn community should adhere to established principles of respect for intellectual property, particularly copyrights when considering the creation of new data sets for training AI models. Avoid uploading confidential and/or proprietary information to AI platforms prior to seeking patent or copyright protection, as doing so could jeopardize IP rights.

University Business Processes

While automating tasks using AI may improve operational efficiency for University Business processes, oversight and review of the use of AI and verification of its outputs for these University business processes should be in place to ensure reliability, consistency, and accuracy.

Additional Guidance

Guidance for Educators

Use of AI should always be in alignment with Penn's [Principles of Responsible Conduct](#). As expectations may vary between classes and instructors, it is important for instructors to provide students with clear guidelines similar to the [guidelines on collaboration](#), on the use of AI within coursework, and when and how the use of AI within a course should be cited. Disclose to students when course materials have been created with the use of AI and when AI detection software will be used in the course.

Be aware of how AI tools may impact assignments and exams in your discipline. [Penn's Center for Teaching and Learning](#) provides useful guidance in this space.

Guidance for Students

All use of AI should be in line with Penn's [Code of Student Conduct](#) and the [Code of Academic Integrity](#).

Educators may have requirements and guidance for citing the use of generative AI output and for attributing AI created content to the specific AI tool and parameters used.

Individual courses may have different or more narrow guidance on the use of AI that should be adhered to within the context of that course.

In the absence of other guidance, treat the use of AI as you would treat assistance from another person. For example, this means if it is unacceptable to have another person substantially complete a task like writing an essay, it is also unacceptable to have AI to complete the task.

Keep in mind that having access to data is not the same as having permission to scrape the data or use it to train an AI model.

Guidance for Researchers

Consult with your department leadership and your discipline's publishing standards to determine how the use of AI should be accounted for with regard to authorship in publications.

Researchers should adhere to federal or international requirements on obtaining informed consent, and Institutional Review Board approvals should be obtained prior to exposing research participant data to AI tools.

Caution should be adopted when research involves the examination of high-risk data, including Personally Identifiable Information (PII) and research participant health information (both identifiable and non-identifiable) exposed to AI.

Intellectual property law may limit what AI can be copyrighted or patented depending on the specifics of their creation, and using data to train AI may implicate copyright laws and/or the patentability of an idea or discovery. Researchers should avoid uploading confidential and/or proprietary information to AI platforms, at least prior to seeking patent or copyright protection, as doing so could jeopardize IP rights. AI users should consult further with the Penn Center for Innovation for additional guidance, including determinations as to whether potential IP rights should be protected prior to the inclusion of data and findings in AI models.

Researchers should reinforce with their mentees (undergraduate students, predoctoral students, and postdoctoral trainees) on the appropriate use of AI.

FAQs AI Guidance

What kinds of AI does this guidance apply to?

This AI Guidance is largely focused on large language models, such as ChatGPT and Google Bard; but these principles apply generally to other machine learning and artificial intelligence technologies.

Who does this AI Guidance apply to?

This Guidance applies to Penn faculty, staff, undergraduate students, pre-doctoral students, and post-doctoral trainees

Who developed this Guidance?

The Offices of Information Security and University Privacy developed this Guidance in collaboration with the Office of the Vice Provost for Education, the Office of the Vice Provost for Faculty, the Office of the Senior Vice Provost for Research, the Office of Clinical Research, the Office of General Counsel, Procurement, and the Penn Center of Innovation.

Can Schools, Centers or departments publish more specific guidance around the use of AI?

Yes, Schools, Centers, and other departments may publish more specific guidance on the use of AI to address certain situations or use cases.

Who do I contact if I have additional questions about this AI Guidance?

For questions about educational policies, please contact the Vice Provost for Education at provost-ed@upenn.edu.

For questions regarding clinical research, please contact the Office of Clinical Research at lauraee@upenn.edu.

For questions about research use, please contact the [Office of the Senior Vice Provost for Research](#) at vpr@upenn.edu.

For questions about procurement, including contracting, sourcing, and supplier management, please contact [Procurement Services](#) at procure@upenn.edu.

For questions regarding intellectual property and AI, contact the Penn Center for Innovation at pciinfo@pci.upenn.edu.

For questions about privacy, please contact the University Privacy Office at privacy@upenn.edu.

For questions about security, please contact the Office of Information Security at security@isc.upenn.edu.

For legal questions related to this guidance document, contact the Office of General Counsel at [Contact Us | OGC \(upenn.edu\)](#).

21. Guidance for the use of generative AI

About this guide

Students and instructors are embracing ChatGPT (GPT-4 as of March 2023) and similar artificial intelligence (AI) technologies across disciplines for different learning goals. There is no one-size-fits-all best practice for their use. This document is meant as a guideline for instructors on what to consider as these tools evolve. We will provide strategies for adopting AI technologies in a responsible, ethical manner, and innovating within each discipline, major, and course. Exploring and communicating about the opportunities and limitations to using these tools will allow instructors and students to critically think about how knowledge is created.

What is ChatGPT?

ChatGPT is a “chatbot” developed by a private company called OpenAI. Users can enter question prompts and within seconds ChatGPT will produce text-based responses in the form of poems,

essays, articles, letters and more. It can also create structured responses like tables, bulleted lists and quizzes. ChatGPT can provide translation and copy language style and structure. It can also be used to develop and debug programming code. New and expanded uses continue to be developed and launched. A similar tool called DALL-E uses AI to create art pieces, and other AI tools have been created or are in rapid development to do even more – music, animation, multimedia video, powerpoints – and the list goes on.

Are students using ChatGPT?

Yes, students have already been exploring and using it to support completion of their coursework. Some courses explicitly encourage the use of ChatGPT for assignments. However, one concern is that students may be using ChatGPT to draft responses to homework responses without learning the material, and this presents the challenges and opportunities for reflection on teaching and learning at this time.

Explore how ChatGPT works

CEILS Education Research Talk with Jess Gregg. Demo portion starts just after the 5 minute mark.
Revolutionizing Education with ChatGPT: How AI is Transforming the Way We Learn – Learning and Technology with Frank

Dig Deeper (article): [How ChatGPT Works](#)

Try out ChatGPT and Reflect: What are the opportunities?

Recommendation

Try out ChatGPT (or other AI tech) with your own course materials and assignments

Instructors can sign up for a free account at [ChatGPT \(or GPT-4\) on Open AI's platform](#). A first step in exploring the tool may be to enter some of your assignment prompts and assess the accuracy

of the output. Then reflect on how you might embrace the tool or implement strategies that make use of the tool unnecessary.

PRO TIP: Including your students in the reflective process is also a learning opportunity to help them understand the benefits and limitations of the tool.

A few things to try:

- **Ask ChatGPT a question** – it could be a homework assignment or any question.
 - How would you evaluate the response provided by ChatGPT?
 - Try modifying the prompt and see how that changes the response.
- **Ask Chat GPT to synthesize text from large documents.** For example, enter a 3500 word paper as a prompt, and ask ChatGPT to create an 18 slide PowerPoint presentation, with headings and bullet points, making a persuasive case for action.
- **Prompt for writing samples specific to your area of expertise.** For example, ask ChatGPT to generate a nurse practitioner note for a 53 year old male with hypertension presenting with shortness of breath and dizziness. Another example is to ask for an email introducing your upcoming course to enrolled students.
- **Ask Chat GPT to translate something.** Together with your students examine the translation to see how well it did. Ask for improvements, or consider when such translation capabilities might be
- **If you teach students how to code,** ask ChatGPT to correct incorrect code (debug code). Consider ways this might help students who are learning to code in your course.
- **If you teach a writing intensive course,** try asking ChatGPT to respond to a writing prompt in a specific style (like a popular author) or create a poem on a specific topic.
 - Reflect on the potential for ChatGPT to support student writing. Which writing skills do you feel are fundamental for students to do independent of artificial intelligence? How might ChatGPT facilitate the development of writing or problem solving skills?
 - Have students use ChatGPT to write a draft and then have them edit what it produces or check for errors

Note: If you do not wish to create an account, reach out to your local teaching support to set up a consultation and explore your assignments together. There are also many demonstrations of ChatGPT available online that you can search and watch.

Ideas for Updating Your Course Activities with AI in Mind

Adjusting assignments and activities

- Consider utilizing ChatGPT and other AI tools explicitly: After experimenting with ChatGPT you may decide that you want to incorporate exercises where students are explicitly encouraged to interact with ChatGPT.

Example:

- Ask students to use ChatGPT and “fact check” the response provided by finding primary and secondary sources to back up the information provided.
- Ask students to generate a first draft using ChatGPT then keep track changes in a document to refine/edit.
- Reflecting upon prompt engineering – use prompting logic used by students to generate information and then provide a different prompt to help guide revision. Showcase that small changes can lead to major differences in output!
- The University of Wisconsin, Madison provides some examples for how to [integrate AI into the writing process in your classroom](#)

View this collaborative [Google Doc: AI Examples and Resources](#) to see examples and resources curated by UCLA’s teaching and learning community.

- Adapt or create assignments that are not easily completed using AI: Be more explicit about having students provide references for assignments, use a social annotation tool like [Hypothes.is](#) or [Perusall](#), utilize comments, Microsoft Word Track Changes or Google Docs Suggesting mode for individual or group annotation, have students complete written assignments in class, ask students to connect learning to their personal experiences and/or current events. Support students in developing oral communication skills by providing more opportunities for in-class presentations (or during discussion sections).

- Montclair State has created a guide that include “[Practical Suggestions to Mitigate Non-Learning/Cheating](#)”
- **Explore Gradescope for implementing additional hand-written assignments:** UCLA has a campus-wide license for Gradescope that [integrates with Bruin Learn](#). One feature of this tool is that students can upload photos or PDFs of written work, and the system easily allows for streamlined grading (by question, by page) and digital commenting and rubrics.
 - You can [watch](#) UCLA’s Will Conley, Department of Mathematics, provide a recorded overview of how to use Gradescope (approx. 56 mins). While we have now shifted to BruinLearn so the initial steps for linking to the gradebook will differ slightly, the within Gradescope interface is the same. [At the 27 minute mark](#) you can view what it looks like to see handwritten work submitted to Gradescope for streamlined digital grading by instructors and TAs.
- **Require students make a connection to class discussions:** Prompt students to explicitly reference in-class discussions, lecture material and course readings in their homework assignments. (Example: Share three takeaways from our in-class discussion on the issues with how Covid-19 testing was implemented across the US at the onset of the pandemic.)

Communicate: Talk With Your Students About AI

Discuss the potential: Many of our students will go on to become leaders at organizations that utilize and/or develop new AI technologies. How will these tools support advancements in your field (medicine, science, art, music, humanities, health, and more)?

Prepare students for the future when they will work and interact with AI: This technology is likely to develop and become embedded in many parts of our lives. Preparing students to thoughtfully engage with it, co-create with it and be curious about and know how to interact with other technological developments as they occur.

Seize the opportunity to center the importance of critical thinking and digital literacy. Students will have the opportunity in the future to break the cycle of spreading disinformation, lack of journalistic integrity in news, and elevating accurate and factual research and scholarship.

Emphasize the importance of digital literacy, research, and writing skills with students; connect students to library resources for research and writing. As educators, we have an obligation to help guide our students through many types of literacy, including digital media and AI literacy. UCLA's [WI+RE](#) has created the [Understanding Misinformation: A Lesson Plan Toolkit](#), that is geared towards educators who want to prepare students to learn about misinformation. We can get students to [vet information like experts](#).

Lean in to talking to your students about how learning happens: Learning happens when actively engaging with the course material, through conversations and dialogue leading to deepening conceptual understanding.

Discuss Academic Integrity with Your Students:

One of the main concerns instructors have expressed is how to uphold academic integrity and prevent the misuse of tools like ChatGPT (intentional or not). These concerns include:

- Plagiarism (copying and pasting the response that the tool provides; running material through multiple AI generators to avoid detection)
- Lack of proper citation of sources
- Inaccurate, misleading, biased, false, or limited information in responses to question prompts. While some AI detectors have been recently developed, it is unclear how effective they will be long-term and early reports indicate that individuals can easily avoid detection through simple modifications to produced text.

The [UCLA Student Conduct Code](#) states, "Unless otherwise specified by the faculty member, all submissions, whether in draft or final form, to meet course requirements (including a paper, project, exam, computer program, oral presentation, or other work) must either be the Student's own work, or must clearly acknowledge the source." Unless an instructor indicates otherwise, the use of ChatGPT or other AI tools for course assignments is akin to receiving assistance from another person and raises the same concern that work is not the student's own. Please communicate this to your students, and consider incorporating this language into your syllabus.

Teaching Assistants will be seeking guidance on how to discuss ChatGPT with students and what to do if they suspect submitted work may be AI generated. In addition to talking with your students, make space for conversations with your TAs and other instructional team members to explore this topic and co-construct guidelines.

The ultimate decision and responsibility for how to teach about AI and the establishment of or revision of course policies related to its use lies with the instructor.

View this collaborative Google Doc: [AI Examples and Resources](#) to see examples (including Syllabus Language and Policy Language) and resources curated by UCLA's teaching and

Discuss the ethical issues and limitations of AI

Facilitate discussions with your students on the impacts of spreading disinformation or biased information, lack of regulation of companies that develop these technologies, and other dangers. While students will likely still continue to use ChatGPT and other tools like it, it is crucial that our community has this shared understanding of both dangers and opportunities.

These technologies are not infallible and their accuracy is subject to a variety of factors, some listed below:

- Prone to filling in replies with incorrect data if there is not enough information available on a subject.
- Lack the ability to understand the context of a particular situation, which can result in inaccurate outputs.
- Large, uncurated datasets scraped from the internet are full of biased data that then informs the models.
- Data is collected from the past, it tends to have a regressive bias that fails to reflect the progress of social movements.

Our nation has yet to catch up to the regulation needed to prevent the potential for tremendous harm when false or biased information is taken as fact. Our community must continue to explore the value and innovation that can come from AI while simultaneously contributing to the dialog about these potential harms.

Share current examples of scholarly discussion on this topic.

As our own UCLA experts have shared with our instructional community during the recent [UCLA Virtual Town Hall: What is ChatGPT and How Does it Relate to UCLA's Academic Mission](#), there are concerns about the ethics and practices around tools like ChatGPT. Algorithms can and do replicate and produce biased, racist, sexist, etc. outputs, along with incorrect and/or misleading information.

Additional Examples:

- Educause – Special Report on Artificial Intelligence
- Noam Chomsky: The False Implications of ChatGPT
- Schools Must Embrace the Looming Disruption of ChatGPT

Be proactive in discussing concerns around privacy and intellectual property that students may have

On Requiring the Use of ChatGPT: Creating an account to use ChatGPT requires sharing of personal information. Depending on context, the use of ChatGPT may also mean sharing student intellectual property or student education records with ChatGPT under their terms and conditions of use. Individual students may have legitimate concerns and therefore may be unwilling to create an account. Discuss these concerns and consider alternatives.

- If you will be requiring use of ChatGPT, consider making this explicit in the syllabus (for a related example of considerations related to privacy, see [Privacy Tips for Your Syllabus](#)).

Protecting student privacy as required by FERPA: Academic records, such as examinations and course assignments, are considered a student record and protected by FERPA. For example, ChatGPT should not be used to draft initial feedback on a student's submitted essay that included their identifying information. Asking ChatGPT to respond to question prompts would not be a FERPA violation, as no student information is provided to ChatGPT.

ChatGPT is currently in the process of review through UCLA's [Third Party Risk Management](#) to understand where there will be gaps in accessibility and security.

Ensure equity and accessibility concerns are addressed

As with any emerging technology, ChatGPT may not always be accessible by individuals with disabilities. Open a conversation with the Center for Accessible Education (CAE) for ideas on exploring accessible alternatives. As always, we encourage faculty to use [this sample syllabus language](#) to direct a student toward CAE to discuss their options for accommodations and support. Requests for support should be directed to caeintake@saonet.ucla.edu or the student's listed Disability Specialist on their accommodation letter.

As the technology evolves, there may be a cost to using it, so continuing to revisit your learning goals and activities with respect to access is a critical equity issue.

Examples and Resources

WEBSITE: [UCLA Online Teaching and Learning – Information on Chat GPT and AI](#) – Resources from UCLA Online Teaching and Learning

CURATED RESOURCE LIST: [Generative AI Tools and Resources](#) – From Dr. Kim DeBacco, Senior Instructional Designer UCLA Online Teaching and Learning

SUBSTACK POST: [“Carving out time to learn: A conversation with ChatGPT”](#) – From Caroline Kong, Instructional Designer and Technologist at the Center for the Advancement of Teaching

RECORDED WEBINAR: [“What’s All the Buzz About ChatGPT and AI in Higher Ed?”](#)

RECORDED WEBINAR: [UCLA Virtual Town Hall: What is ChatGPT and How Does it Relate to UCLA’s Academic Mission](#)

RECORDINGS AND RESOURCES: [AI in Action](#) – Events brought to you by UCLA’s Center for the Advancement of Teaching (CAT), the Center for Education, Innovation, and Learning in the Sciences (CEILS), the Excellence in Pedagogy and Innovative Classrooms program (EPIC), Online Teaching and Learning (OTL), the Bruin Learn Center of Excellence (CoE), the Writing Programs, and Humanities Technology (HumTech)

UCLA NEWSROOM ARTICLE: “Can AI and creativity coexist?” – In a joint interview, UCLA professors Jacob Foster and Danny Snelson discuss how chatbots could be used in teaching, offer historic analogs for the current AI explosion and opine about whether technology is actually capable of creativity.

Example Syllabi Language and Activity Ideas

View this collaborative [Google Doc: AI Examples and Resources](#) to see examples and resources curated by UCLA’s teaching and learning community. This includes syllabus language, assignment ideas, and other strategies shared by instructors from UCLA and across the US.

22. Cornell Guidelines for artificial intelligence

Generative artificial intelligence (AI), offered through tools such as ChatGPT, Claude, Bard, Copilot (Microsoft, GitHub, etc.), and DALL-E, is a subset of AI that uses machine learning models to create new, original content, such as images, text, code, or music, based on patterns and structures learned from existing data.

Guidelines

Cornell’s guidelines seek to balance the exciting new possibilities offered by these tools with awareness of their limitations and the need for rigorous attention to accuracy, intellectual property, security, privacy, and ethical issues. These guidelines are upheld by [existing university policies](#). A key to exploring AI tools centers on important choices about which tools we use and the privacy and protection of an individual’s personal information and institutional data. Free AI tools that are not offered by Cornell do not provide any material protection of data and should not be used to share or process institutional academic or administrative information.

Accountability

You are accountable for your work, regardless of the tools you use to produce it. When using generative AI tools, always verify the information for errors and biases and exercise caution to avoid copyright infringement. Generative AI excels at applying predictions and patterns to create

new content, but since it cannot understand what it produces, the results are sometimes misleading, outdated, or false.

Confidentiality and Privacy

If you are using public generative AI tools, you cannot enter any Cornell information, or another person's information, that is confidential, proprietary, subject to federal or state regulations, or otherwise considered sensitive or restricted. Any information you provide to public generative AI tools is considered public and may be stored and used by anyone else.

As noted in the [University Privacy Statement](#), Cornell strives to honor the [Privacy Principles](#): Notice, Choice, Accountability for Onward Transfer, Security, Data Integrity and Purpose Limitation, Access, and Recourse.

Use for Education and Pedagogy

Cornell encourages a flexible framework in which faculty and instructors can choose to prohibit, to allow with attribution, or to encourage generative AI use. In addition to the [CU Committee Report: Generative Artificial Intelligence for Education and Pedagogy](#) delivered in July 2023 and [resources from the Center for Teaching Innovation](#), check with your college, department, or instructor for specific guidance.

Use for Research

The widespread availability of generative AI tools offers new opportunities of creativity and efficiency and, as with any new tool, depends on humans for responsible and ethical deployment in research and society. The [Cornell University Task Force Report, Generative AI in Academic Research: Perspectives and Cultural Norms](#) (December 2023), offers perspectives and practical guidelines to the Cornell community on the use of generative AI in the practice and dissemination of academic research.

Use for Administration and Other Purposes

By the end of 2023, Cornell is aiming to offer or recommend a set of generative AI tools that will meet the needs of staff doing administrative work, while providing sufficient risk, security, and privacy protections. The use of generative AI for administration purposes must comply with the guidelines of the [Cornell Generative AI in Administration Task Force Report](#) (January 2024).

Cornell AI Services and Pilots

Cornell researchers and other community members have long been involved in AI and machine learning in their scholarly, teaching, and technical work. With the rapid growth of generative AI tools, the Cornell community has more opportunities to enhance our understanding, practice, policy, and technology to use AI in new ways.

Cornell faculty, research, and administrative communities across Ithaca, New York City, and Doha have worked diligently to explore the emerging AI needs, concerns, and potential use cases and lay the groundwork for a deeper understanding of Cornell's near-term institutional needs and potential responses. The resulting university-wide reports are informing AI explorations to meet needs in the teaching and learning, research, and administration spaces.

Throughout spring and summer 2024, we will be piloting a range of services in order to determine where to best invest Cornell's time and talent to meet the university's evolving needs.

For those seeking to purchase generative AI tools or subscriptions outside of those plans, the [IT Statement of Need process](#) is required.

Recommended Resources

[Cornell AI Initiative](#)

[Cornell University AI for Science Institute](#)

[Center for Teaching Innovation: Generative AI](#)

University Task Force Reports

[CU Committee Report: Generative Artificial Intelligence for Education and Pedagogy](#) (June 2023)

[Cornell University Task Force Report - Generative AI in Academic Research: Perspectives and Cultural Norms](#) (December 2023)

[Cornell Generative AI in Administration Task Force Report](#) (January 2024)

AI Questions, Concerns, Needs, or Ideas?

[Contact AI at Cornell](#)

Comments?

To share feedback about this page or request support, [log in](#) with your NetID

23. Generative Artificial Intelligence in the classroom

The latest generation of Artificial Intelligence (AI) systems will impact teaching and learning in many ways, presenting both opportunities and challenges for the ways our course instructors and students engage in learning. At the University of Toronto, we remain committed to providing students with transformative learning experiences and to supporting instructors as they adapt their pedagogy in response to this emerging technology.

While many generative AI systems have recently become available, [ChatGPT](#) is currently the most prominent, garnering worldwide media attention. This is an AI tool that uses predictive technology to create or revise written products of all kinds, including essays, computer code, lesson plans, poems, reports, and letters. The products that the tool creates are generally of good quality, although they can have inaccuracies. We encourage you to try the system to test its capabilities and limitations.

In this FAQ, ChatGPT refers to the free, online AI chat system that utilizes the OpenAI GPT technology. Please note that this is only one of a variety of generative AI tools currently available.

Sample Syllabus Statements

April 2023: The University has created sample statements for instructors to include in course syllabi and course assignments to help shape the message to students about what AI technology is, or is not, allowed. These statements may be used for both graduate and undergraduate level courses.

SGS Guidelines

July 2023: The School of Graduate Studies (SGS) announced new [Guidance on the Appropriate Use of Generative Artificial Intelligence in Graduate Theses](#) which will be of interest to graduate students, supervisors, supervisory committee members, Graduate Chairs and Graduate Units.

Copyright Considerations

September 2023: There remains significant legal uncertainty concerning the use of generative AI tools in regard to copyright. This is an evolving area, and our understanding will develop as new

policies, regulations, and case law become settled. Some of the concerns surrounding generative AI and copyright include:

- Input: The legality of the content used to train AI models is unknown in some cases. There are a number of lawsuits originating from the US that allege Generative AI tools infringe on copyright and it remains unclear if and how the fair use doctrine can be applied. In Canada, there also remains uncertainty regarding the extent to which existing exceptions in the copyright framework, such as fair dealing, apply to this activity.
- Output: Authorship and ownership of works created by AI is unclear. Traditionally, Canadian law has indicated that an author must be a natural person (human) who exercises skill and judgement in the creation of a work. As there are likely to be varying degrees of human input in generated content, it is unclear in Canada how it will be determined who the appropriate author and owner of works are. More recently, the US Copyright Office has published the following guide addressing these issues: [Copyright Registration Guidance for Works Containing AI-Generated Materials](#).

If you have further questions about copyright, please view the U of T Libraries webpage, [Generative AI tools and Copyright Considerations](#) for the latest information.

If you are an instructor who is interested in using generative AI to develop course materials, review the FAQ below for considerations.

Frequently Asked Questions

How can I test out ChatGPT to see its capability?

Instructors are welcome and encouraged to test ChatGPT, use of which is currently free upon registration. You can also test other similar AI tools to assess their capability, for instance to see if they can respond to the assignments used in your courses, or the way in which they improve the readability and grammar of a paragraph. Experimentation is also useful to assess the limits of the tool. However, confidential information should never be entered into an AI tool such as ChatGPT. All content entered may become part of the tool's dataset and may inadvertently resurface in response to other prompts.

Please note that due to high demand, access to ChatGPT is, at times, unavailable.

Is ChatGPT accurate and reliable?

Large Language Models, like ChatGPT, are trained to predict the next word in a sentence, given the text that has already been written. Early attempts at addressing this task (such as the next-word prediction on a smartphone keyboard) are only coherent within a few words, but as the sentence continues, these earlier systems quickly digress. A major innovation of models such as GPT is their ability to pay attention to words and phrases which were written much earlier in the text, allowing them to maintain context for much longer and in a sense remember the topic of conversation. This capacity is combined with a training phase that involves looking at billions of pages of text. As a result, models like ChatGPT, and its underlying technology GPT-3 (and now, GPT-4), are good at predicting what words are most likely to come next in a sentence, which results in generally coherent text.

One area where generative AI tools sometimes struggle is in stating facts or quotations accurately. This means that models like GPT-4 sometimes generate claims that sound real, but to an expert are clearly wrong.

A related area where ChatGPT seems to struggle is in the discussion of any event or concept that has received relatively little attention in online discourse. To assess these limitations, you could try asking the system to generate your biography. Unless there are numerous accurate biographies of yourself online, ChatGPT is unlikely to generate a comprehensively correct biography.

What are the ethical considerations regarding the use of generative AI systems?

This is a threshold question that instructors may want to consider. Mainstream media has been covering this issue extensively, and alternate viewpoints are widely available.

Given that generative AI systems are trained on materials that are available online, it is possible that they will repeat biases present online. OpenAI has invested substantial effort into addressing this problem, but it remains a danger with these types of systems. You may also want to familiarize yourself regarding questions about the way the technology was developed and trained (e.g., who

were the people who trained it?), the way we use the responses it provides, and the long-term impacts of these technologies on the world.

The Provost is consulting with faculty and staff experts on these larger questions involving ChatGPT and other generative AI systems, and welcomes debate and discussion on these issues.

As an instructor, can I use generative AI to generate content for my courses?

The question of copyright ownership remains one of the biggest unknowns when using generative AI tools. The ownership of outputs produced by generative AI is unsettled in law at the current time. If, as an instructor, you would like to use generative AI tools for content generation in your course, consider the following before doing so:

Have an understanding that while you can use these tools to create content, you may not own or hold copyright in the works generated.

Be mindful of what you input into tools: never input confidential information or intellectual property you do not have the rights or permissions to (e.g., do not submit student work or questions without their permission). All content entered may become part of the tool's dataset and may inadvertently resurface in response to other prompts.

Review the terms of service of each tool, which will establish terms of use and ownership of inputs and outputs (for example, view the Terms of Use for OpenAI). Note that terms of use are subject to change without notice.

Be explicit in how you have used these tools in the creation of your work.

View the U of T Libraries, Generative AI tools and Copyright Considerations for more information.

Can I use generative AI tools for pedagogical purposes in my classroom?

Yes. Some instructors may wish to use the technology to demonstrate how it can be used productively, or what its limitations are. The U of T Teaching Centres have developed more information and advice about how you might use generative AI as part of your learning experience design.

24. Generative AI – Academic Integrity at UBC

Thinking about ChatGPT?

Conversations around the impacts of artificial intelligence (AI) tools are ongoing as their capabilities continue to evolve. AI tools have the potential to change the way we teach, learn and work at UBC.

This list brings together important things to know about ChatGPT and generative artificial intelligence in the classroom for instructors and students at UBC. Generative AI technology is evolving quickly and this list will be updated as new developments arise. If you have a question that is not answered here, we invite you to share it through the website [feedback form](#).

What is ChatGPT and what is generative artificial intelligence?

ChatGPT (GPT stands for Generative Pre-trained Transformer) is a tool developed by OpenAI that is capable of producing human-like responses to prompts. This AI system is a large language model that has been trained on a dataset to interact with users in a conversational way. The current version of ChatGPT bases its output on its training data of internet content up to September 2021.

While ChatGPT is widely known, it is part of the broader category of generative AI, a form of artificial intelligence that generates new content based on the data it has been trained on. The new content that is generated can be text, images, code, videos, etc. Other examples of generative artificial intelligence tools (“AI tools”) include Bing, Bard and Dalle-2.

Some generative AI tools are available for free, while others such as ChatGPT Plus have a paid subscription version with further features. Some tools are connected and others offer plugins to connect to the internet and others are not connected and are limited to the information they are trained on which may not be up to the present day.

Can instructors test AI tools to see what they can do, and what they cannot?

Instructors are welcome to test AI tools to assess their capabilities and limitations and experiment with how the tools respond to particular course assignments or prompts. For example, by using prompts from their own assignments and assessments, instructors can gain a sense of how the tool could potentially be used by learners as well as its limitations. Keep in mind that some students may be sophisticated in prompt engineering and could be able to prime GenAI tools to return results that are superior to those obtained by entering an assignment question into a fresh chat.

Additionally, in trying out tools, be sure not to share in prompts any personal or sensitive information, or any information you wish to be kept private, as this content may become part of the dataset the models train on.

Is the use of AI tools considered to be academic misconduct at UBC?

The use of ChatGPT or other generative AI tools does not automatically equate to academic misconduct at UBC. At this time, whether the use of AI tools in courses is or is not allowed is a course or program-level decision and there is no overall, UBC-wide ban on its use in teaching and learning. Individual instructors should clarify expectations with their students early in the term, such as on the syllabus. If instructors have questions about any Department of program level policies on artificial intelligence tools, they should reach out to their Department or program. Further information is available on [generative AI syllabus language](#).

- If using ChatGPT and/or generative AI tools on coursework has been **prohibited** by the instructor, then using these tools would be considered to be academic misconduct.
- If using ChatGPT and/or generative AI tools has been **permitted** by the instructor, then instructors should make sure to convey the limitations of use and how it should be acknowledged and use should stay within those bounds.
- If the use of ChatGPT and/or generative AI tools has **not been discussed or specified** by the instructor, then it is likely to be considered as prohibited as an example of the “use or facilitation of unauthorized means to complete an examination or coursework” and more specifically as “accessing websites or other online resources not specifically permitted by the instructor or examiner” (Discipline for Academic Misconduct, [Vancouver](#) and [Okanagan](#) 3.1.b.iv), and potentially plagiarism (3.1.e).

Students should not assume that all available technologies are permitted. If students are not sure about whether AI tools are allowed, as with any tool, they must ask their instructor for clarity and guidance.

UBC’s Academic Calendar provides guidance on what is considered academic misconduct (Academic Misconduct by UBC Students, [Vancouver](#) and [Okanagan](#)). Academic misconduct is any conduct by which a student gains or attempts to gain an unfair academic advantage or benefit, thereby compromising the integrity of the academic process. Artificial intelligence tools like ChatGPT are not expressly named in the academic misconduct regulation but their use could be

considered an attempt to gain an unfair academic advantage, as well constituting unauthorized means to complete an assignment or assessment, the accessing of a website that is not permitted, or other, depending on the specific case.

It is important for instructors to address academic integrity throughout the semester. Create opportunities to discuss expectations around academic integrity ahead of assignments and exams and throughout the term. Refer to academicintegrity.ubc.ca for **tools to teach academic integrity and respond to academic misconduct.**

How can instructors use AI tools in their classroom?

There are many innovative and creative ways that an instructor might use generative artificial intelligence technologies in their courses. From open conversations around the ethical and societal implications of artificial intelligence to dialogues with chatbots and critical engagement with their outputs, there is potential both for learning about and for learning to use the technology.

If instructors choose to integrate ChatGPT or other GenAI tools into course activities, they should consider the privacy implications of doing so. Using ChatGPT, for example, requires a login that asks for personal information, and it is important to offer students an alternative option if they do not wish to provide this information to engage with the tool themselves. One option could be for the instructor to generate conversations and share them with students to engage with outside of the tool.

Suggestions and examples around how generative AI can be integrated into teaching and learning, as well as how to design assignments and assessments that make it harder to use AI, have been developed by the CTLT. Instructors may wish to review the resource from the CTLT [Teaching and Learning in an Era of Generative AI](#) or [AI in Teaching and Learning](#).

Can students use AI tools to complete assignments?

A key expectation of academic integrity for students is completing their own work. Besides producing essays in seconds, generative AI has proven itself capable of completing multiple-choice exams and short answer questions, generating code, and producing creative output.

While artificial intelligence technologies should not be used to complete academic work, there may be times where instructors and students engage with it as part of student learning. If this is

the case, instructors should provide clear guidance to students around how they are allowed to engage with the tools.

Can or should instructors use AI detectors to detect the use of AI tools on assignments or assessments?

UBC discourages the use of artificial intelligence detectors on student work, and is not at this time planning to purchase or support any such tools at the institutional level.

There are several AI detectors currently in existence, such as GPTZero, Turnitin, and AI Content Detector. Despite the availability of such tools, it is important to remember that they might not be fully tested and that the technology to potentially outwit them continues to evolve. The detectors are not foolproof and can produce false negatives and false positives. It may also be possible for the user to modify content to avoid detection. Instructors might wish to consider [UBC's response](#) to Turnitin's AI detection feature and the concerns that were raised.

If instructors still choose to use AI detectors, they should be aware and understand their limitations and issues they can raise. None of the detectors has undergone a UBC Privacy Impact Assessment, and as such there may be privacy and security concerns with submitting student work to them, particularly without their knowledge or consent. Instructors should not use these tools to evaluate any student work that contains the name of the student or any other personal information of the student or third parties. If student work may be submitted through one or more AI detectors, instructors should be transparent with students and let them know at the beginning of the term, such as in the syllabus.

Finally, it is recommended that AI detection tools not be used as the sole factor in decision-making around an allegation of academic misconduct. If an instructor suspects that an assignment or assessment has been completed with unauthorized use of AI tools, they should proceed as they would for any other potential allegation of academic misconduct. An [overview of the academic misconduct process for instructors](#) is available on the academic integrity website.

Can I use Turnitin's new AI-detection feature in my courses?

No, currently Turnitin's new AI-detection functionality is not available for use in any UBC course.

[UBC has reaffirmed their decision to not enable Turnitin's new AI-detection feature.](#)

On April 4 2023, Turnitin activated **a new and separate feature** that attempts to identify AI-generated text . The LT Hub Leadership group, with the support of the Provosts at both UBC Vancouver and UBC Okanagan, made the decision not to enable this feature at that time ([April 2023](#)) and has recently reaffirmed their decision ([August 2023](#)) for the following reasons:

Effectiveness of the feature is still unclear

- Testing for accuracy in the AI-detection feature is in early stages.
- Testing for potential bias in the feature continues to be in early stages.
- Ability of the feature to keep up with rapidly evolving AI is unknown.

It is not possible to double-check or review the results

- Instructors cannot double-check the feature results.
- Results from the feature are not available for students to review.

You can read more about UBC's **decision to not enable Turnitin's new AI-detection** feature on UBC's Learning Technology Hub website.

UBC is continuing to wait before deploying any AI-detection features, including the one by Turnitin, as more information is needed about the feature's effectiveness, accuracy and bias mitigation, and about the ability to employ an AI detector as a robust component of talking to students about suspected academic misconduct. The use of other AI detection tools is also not recommended, due to similar issues as those noted above, as well as privacy and security concerns; to date, no AI detection tool has undergone a UBC Privacy Impact Assessment process.

What should I do if I suspect the unauthorized use of AI tools?

Instructors who suspect that a student has used generative artificial intelligence tools contrary to expectations should follow the **standard academic misconduct process**. If an instructor has a suspicion based on the student's work, they should follow the procedure as they would for any misconduct allegation. Instructors should not rely on AI detectors to form the basis of an allegation of academic misconduct . If students have concerns about any allegations of academic misconduct against them, they can reach out to the **Office of the Ombudsperson for Students, AMS Advocacy or SUO Advocacy**.

What are the recommendations for citing content developed by Generative AI if its use is permitted?

There has been much conversation about how generative AI should be cited if its use is permitted in assignments and academic publications. The [American Psychological Association](#) (APA), [Modern Language Association](#) (MLA) and the [Chicago Manual of Style](#) have all provided recommendations in this area. Further information on this topic can be found in the [Generative AI and ChatGPT LibGuide](#) from the UBC Library.

Is there funding available for teaching and learning projects on generative AI?

Funding is available for projects that explore the intersection of generative artificial intelligence and teaching and learning, as well as other teaching and learning topics ([Teaching and Learning Enhancement Fund](#), [Aspire 2040 Learning Transformations Fund](#), [SOTL Seed Program](#)). Explore teaching and learning topics and further support through the [CTL](#) and the [CTLT](#).

25. Principles on Generative AI in Teaching and Learning at McGill

Generative Artificial Intelligence (AI) tools have broad applications across society, including for institutions of higher education. The Subcommittee on Teaching and Learning (STL) of McGill's Academic Policy Committee (APC) established a working group in January of 2023 to develop an approach regarding Generative AI and its impacts on Teaching and Learning. Their final report, available online, was discussed at APC on 26 October 2023, and the recommendations were endorsed by McGill's Senate on 15 November 2023. The five principles emerging from the report are reproduced below. These provide a framework for ongoing conversations about Generative AI at the University in the context of teaching and learning, and can be used as a guide for instructors, students, and staff, and also for Faculties when they consider their own internal guidelines. Some of the principles contain within them areas of ongoing focus and attention, notably around the importance of education and awareness. Members of the community can also consult the Teaching and Learning Knowledge base for additional information about Generative AI

26. AI-Squared – Artificial Intelligence and Academic Integrity

At the outset, a number of universities around the world elected to ban the student use of ChatGPT (as well as other Generative AI Tools) in academic work based on the news that ChatGPT was capable of generating outputs of unexpected quality and sophistication. Academic integrity concerns, both real and imagined, were raised over how students might use ChatGPT inappropriately in their academic assessments (Shiri 2023). The majority of these concerns are

premised on the view that Generative AI-tools have the potential to do more harm than good, and, that, when wielded by uninformed students, the technology's impact ranges from quick 'work-around' to weapon (Sawahel 2023, Weisman 2023).

To AI or Not to AI?

While some individuals and institutions may view Generative AI with suspicion and as a threat to higher education, there are others who want to weigh the options from more nuanced perspectives. Many universities, for example, have put together task forces to make teaching in the context of AI a priority (Baucom, nd.; Black, 2023; Fox, 2023). When used ethically and in pedagogically sound ways, AI-tools can offer academics the chance to reconsider and reimagine an educational focus, not on deliverables and summative end-products (such as written assignments and standard exams) as measures of learning, but instead on process-driven and evaluated assessment. Stated another way, learning is not only about the product; learning is also about the process of acquiring new knowledge or learning ways to think and reason. This gives the instructor a window through which to focus on what students are 'doing' in their classes to develop the requisite disciplinary knowledge and allied critical-thinking abilities. While tools like ChatGPT are prone to fabrication (factually inaccurate outputs) and generating biases found in its data lakes, they can also help to deepen student engagement and enhance teaching, learning, and assessment (Mollick, 2023). Importantly, these technologies should be recognized as potential tools enabling increased accessibility to learning that will support a wider diversity of student needs than previously possible.

Therein lies another important consideration: AI and its various tools are well on their way to becoming omnipresent in our lives. Learning to adapt to AIs presence in our academic spaces is part of teaching today's learners. To simply ban or avoid AI is to avoid the reality that U of A students are engaging and experimenting with (i.e. yes, like right now, as you read this) with and experiment with AI tools in our courses.

In order to do so, instructors are encouraged to experiment with different approaches, and to find ways to adapt and improve teaching and assessment to embrace the new reality of working and studying in a world where these emerging technologies are freely and widely available.

Guidance - University of Alberta's Provost's Task Force on Artificial Intelligence and the Learning Environment

The suggestions provided in this section are in alignment with the initial [Guidance](#) proposed by the University of Alberta's Provost's Task Force on Artificial Intelligence and the Learning Environment (March 2, 2023).

Given the diversity of learning environments across our campuses, the general guidance that we can give includes the following:

1. Have conversations with your students about your expectations regarding the use of Generative AI, particularly in your course assignments. If students are using Generative AI, how would you like them to indicate that to you (e.g. in the sources cited page, methodology section, prefatory comments, or in-text citation)? Please make sure that you also summarize these conversations in a written format and include them in eClass in a place where students will find them for those who may not have been in class. This also gives students a place to refer back to when completing assignments. Your Department or Faculty may also have specific guidance for you.
2. Identify creative uses for Generative AI in your course (idea generation; code samples; creative application of course concepts; study assistance; language practice). Discuss the limitations of tools like ChatGPT in the topics covered by your course, including the limitation of data used (prior to 2021), factually inaccurate information, biases and discrimination in the data used to generate text and in the output, and the use of culturally inappropriate language and sources.
3. Remind students that the Code of Student Behaviour states: “No Student shall represent another’s substantial editorial or compositional assistance on an assignment as the Student’s own work.” Submitting work created by generative AI and not indicating such would constitute cheating as defined above.
4. Stress to students the value of building their own voice, writing skills, and so on. Motivating students to share their ideas, perspectives, and voice may make generative AI less appealing. Similarly, asking students to share their reflections (reflective writing) can help reinforce student investment in the learning process. If instructors are equipped to do so, they can even show how generative AI can be used as a tool to aid in work as opposed to replacing student work.
5. Remind students that AI tools such as ChatGPT gather significant personal data from users to share with third parties.

In order to help you think through the various options available to you, CTL suggests you start with the following questions as part of your decision-making process:

1. What are your discipline's conventions and assumptions? How might students use AI to support their academic work in your discipline?
2. What role, if any, do AI driven technologies in the course/classroom play in your personal teaching philosophy?
3. Is assessment task redesign needed? How significant is this redesign and development? How do the new assessments fit and align with the course learning outcomes?
4. What do you want your students to know about your expectations regarding AI and academic integrity?
5. Which University resources would you like to direct your students to for further guidance if necessary?
6. What kind of classroom environment would your students like to see? How might you include them in the conversation about AI use in academic work?

Where did you land with your responses to the above questions? Are you leaning towards experimenting with/integrating AI tools in your teaching? Or, do you plan not to use AI in your classroom but still allow students to use it for specific purposes in the course of their learning? Whatever your final decision, it is important that you be transparent and share this information with your students.

If you are a YES to AI use:

Stress to your students that if they use generative AI in their academic work, it is important they do so honestly, transparently, and according to the expectations you set for them. The substance of these conversations should match the language of expectations spelled out in any reference documentation you provide to students, such as the course syllabus or a [Statement of Expectations for AI Use](#). Another issue for you to consider, it is very possible that you may encounter students in your courses who do not want to use AI tools. In such cases – and assuming AI is not integral to course contents and learning outcomes – instructors can offer alternative assessment tasks.

If you are a NO to AI use:

Let students know that although the University of Alberta's most recent (November 2022) [Code of Student Behaviour](#) does not explicitly reference Generative AI and its use, if a student submits academic work, including text, images, code, and designs, generated by AI without proper

attribution, instructors can consider this an act of plagiarism under the Code, which states: “No Student shall represent another’s substantial editorial or compositional assistance on an assignment as the Student’s own work. If you are a NO, it is important for you to consider that not every use of AI tools by students may qualify as cheating – students might use the tools in ways that support and deepen their learning.

Tenets of Postplagiarism: Writing in the Age of Artificial Intelligence

Responding to concerns about AI and writing, plagiarism, and academic integrity, Sarah E. Eaton contemplates a future in which we enter an era marked for “postplagiarism.” This will be a time when archaic print-based notions of copyright are set aside to make way for human-AI partnerships and new definitions of authorship and originality.

Although Eaton first put forward these ideas in her book, [*Plagiarism in Higher Education: Tackling Tough Topics in Academic Integrity \(2021\)*](#), she has recently revised her thinking and put forward [*6 Tenets of Postplagiarism: Writing in the Age of Artificial Intelligence*](#).

Eaton identifies six notions that will likely come to characterize the age of postplagiarism:

1. Hybrid Human-AI Writing Will Become Normal
2. Human Creativity is Enhanced
3. Language Barriers Disappear
4. Humans can Relinquish Control, but not Responsibility
5. Attribution Remains Important
6. Historical Definitions of Plagiarism No Longer Apply

For more information, please also see her article, [*“Artificial intelligence and academic integrity, post-plagiarism” \(2023\)*](#).

Dialogue with students

Begin with a conversation, in-person or synchronously, if possible, so you have the best opportunity to openly dialogue about your expectations and gauge your students’ responses. Talk to your students. The purpose of this initial dialogue is to share with them your expectations, and explore together in two-way conversation the possibilities and limitations of using Generative AI tools in the context of your course(s) and their academic work in your discipline. Speak to them about the academic integrity concerns that have been raised at the U of A and elsewhere in higher

education. Where appropriate, encourage your students to ask questions, provide inputs, and offer suggestions. You might be surprised to discover instructors sometimes need to explore AI basics with their students too. Not all students will be up to date on the AI.

A few key questions to guide your conversation with your students include:

1. What do you know about artificial intelligence and AI tools such as GPT-4, Midjourney, and Microsoft's (GPT-powered) search engine, Bing?
2. Have you used any of them before? Why?
3. Have you used an AI tool for learning (specifically, in your academic work)?
4. If so, how did you use them?
5. How do you think you can ethically use AI tools to support your learning?

(Adapted from, Eaton, 2023)

This conversation is a great opportunity for you to discuss (November 2022) [Code of Student Behaviour](#) and [Academic Misconduct](#) with students, so, together, you can all consider the ethical implications and responsibilities.

If you plan to integrate AI into your in-person or hybrid courses, here are a couple of options you can use to continue the conversation:

1. Create an AI-based Discussion forum to share in (and monitor) your students' experiences and conversations about their use of AI tools.
2. Create a Journal activity, and request that your students transparently track and reflect on their use of AI-tools as part of their learning process during your class.

27. Artificial Intelligence and ChatGPT – Academic Integrity

Quick Links

- [Instructor resources for Artificial Intelligence & ChatGPT tools](#)
- [Tips for students](#)

Artificial intelligence is everywhere and with it comes new opportunities and challenges.

As these new platforms evolve so must the way instructors, TAs and students use them. Given the rapid pace of change and growth in artificial intelligence - this issue will be ongoing for the foreseeable future.

At present, **it is important for instructors to be explicit about whether artificial intelligence or tools like ChatGPT are allowed to be used to complete assignments, tests or exams, and if so, the extent to which it is allowed, and if it should be cited and how to cite it**. A student who does not comply with the instructors rules about the use of such tools will be subject to Policy 71 and an investigation into academic misconduct.

A memo from the Associate Vice-President, Academic about Chat GPT (and other AI tools) can be found [here](#).

FAQ for Instructors and TAs can be found [here](#). Wiley has also put out a publication on *AI in Higher Education: Current Uses and Future Applications* which can be found [here](#).

Thinking about Generative Artificial Intelligence in the Classroom (Video): An overview for University of Waterloo instructors on considerations for generative artificial intelligence in teaching and learning. Includes recommendations instructors may wish to implement in their teaching practice.

INSTRUCTOR RESOURCES FOR ARTIFICIAL INTELLIGENCE & CHATGPT TOOLS

- [AIO Guidance Regarding Student Use of AI in Your Course](#) (Toronto Metropolitan University)
- [Teaching and Learning with Artificial Intelligence Apps](#) (University of Calgary)
- [Artificial intelligence and Assessment Design](#) (Flinders University)
- [Artificial Intelligence in Teaching & Learning](#) (University of Toronto)
- [Conversations with Students about Generative Artificial Intelligence Tools](#) (Centre for Teaching Excellence)

TIPS FOR STUDENTS

Article: [5 things students need to know before using GenAI](#)

Interactive Lesson - [The Assignment Dilemma: Using GenAI Productively](#)

This interactive lesson offers an engaging experience that delves into the ethical considerations, challenges, and applications of Generative Artificial Intelligence (AI). By participating in this interactive lesson, you can better understand the practical aspects of handling Generative AI responsibly.

Before beginning a project/assignment/test:

- Review the instructions and ensure you are clear on your instructors expectations. If you are unsure if you can use AI tools, ask your instructor **before** starting the project/assignment/test.
- If AI tools are permitted in a project or written assignment, ensure you are aware of how to cite the use of such tools. If you are not sure which citation style to use or how to cite AI tools, ask your instructor. A helpful guide to citing AI tools in APA style can be found [here](#).
- Check out resources from the Writing and Communication Centre [here](#) that instruct you how to use GenAI tools in the writing process.
- Develop a plan working backwards from the due date.
- Create a schedule to stay on track. Find schedule templates in the SSO's section on '[Time Management](#)'.
- If you find you are falling behind and might not make the due date reach out to your instructor immediately.
- Reach out to campus resources like the [Writing and Communication Centre](#) or [SSO](#) for support and guidance.

The University of Queensland Australia has developed a brief module for students on ways to incorporate AI tools into their studies. To review the module click [here](#).

Remember you only get out what you put into your academic courses! To graduate with the best skills for your future - ensure you are learning the content and can demonstrate your knowledge.

28. Montreal Declaration on Responsible AI

On November 3, 2017, the Université de Montréal launched the co-construction process for the Montréal Declaration for a Responsible Development of Artificial Intelligence (Montréal Declaration). A year later, the results of these citizen deliberations are public. Dozens of events were organized to stimulate discussion on social issues that arise with artificial intelligence (AI), and 15 deliberation workshops were held over three months, involving over 500 citizens, experts and stakeholders from all backgrounds.

The Montréal Declaration is a collective work that aims to put AI development at the service of the well-being of all people, and to guide social change by developing recommendations with strong democratic legitimacy.

The selected citizen co-construction method is based on a preliminary declaration of general ethical principles structured around seven (7) fundamental values: well-being, autonomy, justice, privacy, knowledge, democracy and responsibility. Following the process, the Declaration was enriched and now presents 10 principles based on the following values: well-being, autonomy, intimacy and privacy, solidarity, democracy, equity, inclusion, caution, responsibility and environmental sustainability.

Our process for responsible artificial intelligence

The Montreal Declaration for a Responsible Development of Artificial Intelligence is based on a declaration of ethical principles built around 7 core values: well-being, autonomy, justice, privacy, knowledge, democracy and responsibility. These values, suggested by a group of ethics, law, public policy and artificial intelligence experts, have been informed by a deliberation process. This deliberation occurred through

consultations held over three months, in **15 different public spaces**, and sparked exchanges between **over 500 citizens, experts and stakeholders** from every horizon.

Educational, ethical and methodological introduction

What is artificial intelligence? What are the ethical issues raised by AI? What is co-construction, and most importantly, what is expected of citizens? The scientific co-directors of the Declaration answer these questions and set the stage for the ensuing discussions.

Debating

Education, Health, Smart Cities, Justice and the Workforce are the 5 sectors around which prospective scenarios were developed. Using these scenarios set in 2025, groups of 5 to 8 people, with the help of a facilitator, discuss ethical issues.

Suggesting

Using the issues developed for 2025, participants must now imagine recommendations to allow a responsible rollout and use of AI in Quebec.

29. Provisional Guidelines on the Use of Generative AI in Teaching and Learning

Provisional Principles

These overarching provisional principles have guided the work of the Task Force on Generative AI in Teaching and Learning.

- Students want to learn, and instructors want to support their learning.
- Participatory learning – learning which happens in relationships and community – continues to be a valuable and vital way for students to learn.

- Assessments that require students to document the process of learning continue to be meaningful for student learning.
- Generative AI poses risks, as well as opportunities. Individuals will have different reactions and different expectations for the technology.

Provisional Guidelines on the Use of Generative AI in Teaching and Learning

General Guidelines

Academic Integrity

McMaster's existing [academic integrity policy](#) applies when using generative AI. Its overall definition of academic dishonesty, which is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage, allows for allegations related to generative AI. The policy states under item 18(c) that "It shall be an offence knowingly to ... submit academic work for assessment that was purchased or acquired from another source".

- a. Unless otherwise stated, students should assume use of generative AI is prohibited.
 - b. Instructors who incorporate generative AI into courses should explain to students in writing and verbally in-class how generative AI material should be acknowledged or cited. [See McMaster's citation guide](#) for examples.
2. Generative AI plagiarism detection software is currently unavailable or not recommended at McMaster. This software will continue to be reviewed and may be used in the future.
 - a. These detectors will produce false positives and are not approved for use through the University's policy. Students have not consented to the sharing of their intellectual work through these tools. It is also unclear how the material submitted to the third-party detectors is retained or used.
 - b. Until more is understood about generative AI detection tools, instructors should not submit student work to generative AI detection tools.
 - c. McMaster has an institutional membership to Turnitin, a plagiarism detection software. Turnitin [announced an update](#) aimed at detecting writing produced by generative AI. McMaster, like many other institutions, has not yet turned on this feature as there is a need to understand the functionality of the tool, assess the

security and privacy considerations for student work and determine whether it aligns with existing policies.

- d. If you do suspect student work may have violated the academic integrity policy, please review the steps to take.

Instructor or TA Use of Generative AI

1. If instructors use generative AI in their teaching materials instructors should explain in the course outline the extent to which generative AI has been, or will be, used.
 - a. Instructors should fact-check any generative AI produced materials.
 - b. Instructors should not submit student work to generative AI tools for feedback without students' consent and ability to opt-out.
2. Course instructors have three options for directing teaching assistant use of generative AI
 - a. Permitting teaching assistants to use generative AI for any aspect of teaching assistant work, with the exception of summative evaluation, with no expectation that they use generative AI and no training specific to generative AI required. TAs must inform the instructor of the intended use of generative AI, and receive approval, before implementation. Summative evaluations are those which significantly impact a student's grade or progress in a course. This includes providing a quantitative grade (number or letter grade).
 - b. Requiring teaching assistants to use generative AI for specified teaching tasks as outlined in the hours of work form and with training provided.
 - In the instance of required use: As directed by the course instructor explicitly in the hours of work form, teaching assistants will use generative AI for the specific teaching tasks. Course instructors will provide teaching assistants with the necessary training to use generative AI for the specified teaching purpose(s) with this training included in the hours of work. Teaching assistants will evaluate all teaching materials/formative feedback developed with generative AI for accuracy before use with students. Any planned use of generative AI by teaching assistants will be shared with students in the course outline.
 - c. Prohibiting teaching assistants from using generative AI for teaching tasks

3. Generative AI tools can be used to provide formative feedback on student work; generative AI tools cannot be used to provide summative evaluation of student work.

- a. AI-generated formative feedback is intended to guide learning and improve understanding, by pointing out strengths and areas for improvement in student work.
- b. Summative evaluations are those which significantly impact a student's grade or progress in a course. This includes providing a quantitative grade (number or letter grade).
- c. Instructors, or teaching assistants when directed, should review AI-generated formative feedback to ensure it aligns with the learning objectives and course materials, and add their own insights where necessary. Formative feedback that uses AI should not be given a quantitative grade by the AI tool. A "pass/fail" or "completion" may be applied.
- d. Instructors, or teaching assistants when directed, are responsible for summative evaluations to ensure appropriateness and accuracy.
- e. Data collection should be turned off on generative AI tools when used for providing formative feedback.
- f. When providing AI-generated formative feedback, students should be made aware that it is generated by AI explicitly in the course syllabus.

4. Instructors who include assessments that incorporate generative AI should:

- a. Consider including reflective components that invite students to comment on the use of/experience with generative AI in the assessment
- b. Explicitly review criteria and/or rubrics in ways that demonstrate how the use of generative AI is being assessed.

Privacy, Security and Selection of Tools

- 1. Instructors incorporating generative AI should be aware of the privacy policies and user agreements of each generative AI tool and alert students to these policies in the course outline.
- 2. Where possible, courses that incorporate generative AI should rely on free versions of generative AI tools (e.g. Microsoft Copilot, ChatGPT 3) for student use.

- a. Alternatives should be provided for Generative AI tools that are restricted to users 18+ (e.g. ChatGPT).

Student Assessments

1. Assessment alternatives that may be less susceptible to the use of generative AI include oral exams, presentations followed by a Q and A, invigilated/in-class assessments, practical tests, assessments that incorporate class discussion/activities, and process-based work.
2. Instructors may consider adding an honour pledge to assessments.
3. Students may opt-out of assessments that require the use of generative AI only in exceptional circumstances as approved by the course instructor. If approved to opt-out of an assessment that requires the use of generative AI based on an exceptional circumstance, students will not face academic penalty, but will be required to provide alternative and equivalent evidence of their learning as proposed to, and agreed to by, the course instructor.

Ongoing Support and Continued Work

1. The MacPherson Institute will continue to provide training and resources for instructors and students on how to use generative AI effectively. See [the MacPherson Institute website](#) for current workshops, resources and to schedule a consultation.
2. McMaster will explore an annual donation to carbon offsetting programs to address the environmental impact of training large AI models.
3. The MacPherson Institute will collect feedback from instructors and students this fall on their experiences, questions and concerns about using generative AI in teaching and learning in an effort to update and improve these guidelines.
4. These guidelines will be regularly reviewed and revised.

Sample Syllabus Statements

These sample syllabus statements may be included on a course syllabus to communicate with students the expectations around generative AI in a course. Instructors may adapt or modify these statements according to their individual teaching goals and course learning outcomes.

Use Prohibited

Students are not permitted to use generative AI in this course. In alignment with [McMaster academic integrity policy](#), it “shall be an offence knowingly to ... submit academic work for assessment that was purchased or acquired from another source”. This includes work created by generative AI tools. Also state in the policy is the following, “Contract Cheating is the act of “outsourcing of student work to third parties” (Lancaster & Clarke, 2016, p. 639) with or without payment.” Using Generative AI tools is a form of contract cheating. Charges of academic dishonesty will be brought forward to the Office of Academic Integrity.

Some Use Permitted

Example One

Students may use generative AI in this course in accordance with the guidelines outlined for each assessment, and so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside assessment guidelines or without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the limitations for use for each assessment and to be clear on the expectations for citation and reference and to do so appropriately.

Example Two

Students may use generative AI for [editing/translating/outlining/brainstorming/revising/etc] their work throughout the course so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside the stated use of [editing/translating/outlining/brainstorming/revising/etc] without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the limitations for use and to be clear on the expectations for citation and reference and to do so appropriately.

Example Three

Students may freely use generative AI in this course so long as the use of generative AI is referenced and cited following citation instructions given in the syllabus. Use of generative AI outside assessment guidelines or without citation will constitute academic dishonesty. It is the student’s responsibility to be clear on the expectations for citation and reference and to do so appropriately.

Unrestricted Use

Students may use generative AI throughout this course in whatever way enhances their learning; no special documentation or citation is required.

30. Artificial Intelligence and the Future of Teaching and Learning

Introduction

The U.S. Department of Education (Department) is committed to supporting the use of technology to improve teaching and learning and to support innovation throughout educational systems. This report addresses the clear need for sharing knowledge and developing policies for “Artificial Intelligence,” a rapidly advancing class of foundational capabilities which are increasingly embedded in all types of educational technology systems and are also available to the public. We will consider “educational technology” (edtech) to include both (a) technologies specifically designed for educational use, as well as (b) general technologies that are widely used in educational settings. Recommendations in this report seek to engage teachers, educational leaders, policy makers, researchers, and educational technology innovators and providers as they work together on pressing policy issues that arise as Artificial Intelligence (AI) is used in education. AI can be defined as “automation based on associations.” When computers automate reasoning based on associations in data (or associations deduced from expert knowledge), two shifts fundamental to AI occur and shift computing beyond conventional edtech: (1) from capturing data to detecting patterns in data and (2) from providing access to instructional resources to automating decisions about instruction and other educational processes. Detecting patterns and automating decisions are leaps in the level of responsibilities that can be delegated to a computer system. The process of developing an AI system may lead to bias in how patterns are detected and unfairness in how decisions are automated. Thus, educational systems must govern their use of AI systems. This report describes opportunities for using AI to improve education, recognizes challenges that will arise, and develops recommendations to guide further policy development.

Rising Interest in AI in Education

Today, many priorities for improvements to teaching and learning are unmet. Educators seek technology-enhanced approaches addressing these priorities that would be safe, effective, and scalable. Naturally, educators wonder if the rapid advances in technology in everyday lives could help. Like all of us, educators use AI-powered services in their everyday lives, such as voice assistants in their homes; tools that can correct grammar, complete sentences, and write essays; and automated trip planning on their phones. Many educators are actively exploring AI tools as they are newly released to the public¹. Educators see opportunities to use AI-powered capabilities like speech recognition to increase the support available to students with disabilities, multilingual learners, and others who could benefit from greater adaptivity and personalization in digital tools for learning. They are exploring how AI can enable writing or improving lessons, as well as their process for finding, choosing, and adapting material for use in their lessons. Educators are also aware of new risks. Useful, powerful functionality can also be accompanied with new data privacy and security risks. Educators recognize that AI can automatically produce output that is inappropriate or wrong. They are wary that the associations or automations created by AI may amplify unwanted biases. They have noted new ways in which students may represent others' work as their own. They are well-aware of "teachable moments" and pedagogical strategies that a human teacher can address but are undetected or misunderstood by AI models. They worry whether recommendations suggested by an algorithm would be fair. Educators' concerns are manifold. Everyone in education has a responsibility to harness the good to serve educational priorities while also protecting against the dangers that may arise as a result of AI being integrated in edtech. To develop guidance for edtech, the Department works closely with educational constituents. These constituents include educational leaders—teachers, faculty, support staff, and other educators—researchers; policymakers; advocates and funders; technology developers; community members and organizations; and, above all, learners and their families/caregivers. Recently, through its activities with constituents, the Department noticed a sharp rise in interest and concern about AI. For example, a 2021 field scan found that developers of all kinds of technology systems—for student information, classroom instruction, school logistics, parent teacher communication, and more—expect to add AI capabilities to their systems. Through a series of four listening sessions conducted in June and August 2022 and attended by more than 700 attendees, it became clear that constituents believe that action is required now in order to get ahead of the expected increase of

AI in education technology—and they want to roll up their sleeves and start working together. In late 2022 and early 2023, the public became aware of new generative AI chatbots and began to explore how AI could be used to write essays, create lesson plans, produce images, create personalized assignments for students, and more. From public expression in social media, at conferences, and in news media, the Department learned more about risks and benefits of AI-enabled chatbots. And yet this report will not focus on a specific AI tool, service, or announcement, because AI-enabled systems evolve rapidly. Finally, the Department engaged the educational policy expertise available internally and in its relationships with AI policy experts to shape the findings and recommendations in this report.

Three Reasons to Address AI in Education Now

“I strongly believe in the need for stakeholders to understand the cyclical effects of AI and education. By understanding how different activities accrue, we have the ability to support virtuous cycles. Otherwise, we will likely allow vicious cycles to perpetuate.” —Lydia Liu During the listening sessions, constituents articulated three reasons to address AI now: First, AI may enable achieving educational priorities in better ways, at scale, and with lower costs. Addressing varied unfinished learning of students due to the pandemic is a policy priority, and AI may improve the adaptivity of learning resources to students’ strengths and needs. Improving teaching jobs is a priority, and via automated assistants or other tools, AI may provide teachers greater support. AI may also enable teachers to extend the support they offer to individual students when they run out of time. Developing resources that are responsive to the knowledge and experiences students bring to their learning—their community and cultural assets—is a priority, and AI may enable greater customizability of curricular resources to meet local needs.

As seen in voice assistants, mapping tools, shopping recommendations, essay-writing capabilities, and other familiar applications, AI may enhance educational services. Second, urgency and importance arise through awareness of system-level risks and anxiety about potential future risks. For example, students may become subject to greater surveillance. Some teachers worry that they may be replaced—to the contrary, the Department firmly rejects the idea that AI could replace teachers. Examples of discrimination from algorithmic bias are on the public’s mind, such as a voice recognition system that doesn’t work as well with regional dialects, or an exam monitoring system that may unfairly identify some groups of students for disciplinary action. Some uses of AI may be infrastructural and invisible, which creates concerns about transparency and trust. AI often

arrives in new applications with the aura of magic, but educators and procurement policies require that edtech show efficacy. AI may provide information that appears authentic, but actually is inaccurate or lacking a basis in reality. Of the highest importance, AI brings new risks in addition to the well-known data privacy and data security risks, such as the risk of scaling pattern detectors and automations that result in “algorithmic discrimination” (e.g., systematic unfairness in the learning opportunities or resources recommended to some populations of students). Third, urgency arises because of the scale of possible unintended or unexpected consequences. When AI enables instructional decisions to be automated at scale, educators may discover unwanted consequences. In a simple example, if AI adapts by speeding curricular pace for some students and by slowing the pace for other students (based on incomplete data, poor theories, or biased assumptions about learning), achievement gaps could widen. In some cases, the quality of available data may produce unexpected results. For example, an AI-enabled teacher hiring system might be assumed to be more objective than human-based résumé scoring. Yet, if the AI system relies on poor quality historical data, it might de-prioritize candidates who could bring both diversity and talent to a school’s teaching workforce. In summary, it is imperative to address AI in education now to realize key opportunities, prevent and mitigate emergent risks, and tackle unintended consequences.

Toward Policies for AI in Education

The 2023 AI Index Report from the Stanford Institute for Human-Centered AI has documented notable acceleration of investment in AI as well as an increase of research on ethics, including issues of fairness and transparency.² Of course, research on topics like ethics is increasing because problems are observed. Ethical problems will occur in education, too.³ The report found a striking interest in 25 countries in the number of legislative proposals that specifically include AI. In the United States, multiple executive orders are focused on ensuring AI is trustworthy and equitable, and the White House Office of Science and Technology Policy has introduced a Blueprint for an AI Bill of Rights (Blueprint)⁴ that provides principles and practices that help achieve this goal. These initiatives, along with other AI-related policy activities occurring in both the executive and legislative branches, will guide the use of AI throughout all sectors of society. In Europe, the European Commission recently released Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators.⁵ AI is moving fast and heralding societal changes that require a national policy response. In addition to broad policies for all sectors of society, education-specific policies are needed to address new opportunities and challenges within existing

frameworks that take into consideration federal student privacy laws (such as the Family Educational Rights and Privacy Act, or FERPA), as well as similar state related laws. AI also makes recommendations and takes actions automatically in support of student learning, and thus educators will need to consider how such recommendations and actions can comply with laws such as the Individuals with Disabilities Education Act (IDEA). We discuss specific policies in the concluding section.

AI is advancing exponentially (see Figure 1), with powerful new AI features for generating images and text becoming available to the public, and leading to changes in how people create text and images⁶. The advances in AI are not only happening in research labs but also are making news in mainstream media and in educational-specific publications. Researchers have articulated a range of concepts and frameworks for ethical AI⁷, as well as for related concepts such as equitable, responsible, and human-centered AI. Listening session participants called for building on these concepts and frameworks but also recognized the need to do more; participants noted a pressing need for guardrails and guidelines that make educational use of AI advances safe, especially given this accelerating pace of incorporation of AI into mainstream technologies. As policy development takes time, policy makers and educational constituents together need to start now to specify the requirements, disclosures, regulations, and other structures that can shape a positive and safe future for all constituents—especially students and teachers. Policies are urgently needed to implement the following: 1. leverage automation to advance learning outcomes while protecting human decision making and judgment; 2. interrogate the underlying data quality in AI models to ensure fair and unbiased pattern recognition and decision making in educational applications, based on accurate information appropriate to the pedagogical situation; 3. enable examination of how particular AI technologies, as part of larger edtech or educational systems, may increase or undermine equity for students; and 4. take steps to safeguard and advance equity, including providing for human checks and balances and limiting any AI systems and tools that undermine equity.

Building Ethical, Equitable Policies Together

In this report, we aim to build on the listening sessions the Department hosted to engage and inform all constituents involved in making educational decisions so they can prepare for and make better decisions about the role of AI in teaching and learning. AI is a complex and broad topic, and we are not able to cover everything nor resolve issues that still require more constituent engagement.

This report is intended to be a starting point. The opportunities and issues of AI in education are equally important in K-12, higher education, and workforce learning. Due to scope limitations, the examples in this report will focus on K-12 education. The implications are similar at all levels of education, and the Department intends further activities in 2023 to engage constituents beyond K-12 schools.

Guiding Questions

Understanding that AI increases automation and allows machines to do some tasks that only people did in the past leads us to a pair of bold, overarching questions: 1. What is our collective vision of a desirable and achievable educational system that leverages automation to advance learning while protecting and centering human agency? 2. How and on what timeline will we be ready with necessary guidelines and guardrails, as well as convincing evidence of positive impacts, so that constituents can ethically and equitably implement this vision widely? In the Learning, Teaching, and Assessment sections of this report, we elaborate on elements of an educational vision grounded in what today's learners, teachers, and educational systems need, and we describe key insights and next steps required. Below, we articulate four key foundations for framing these themes. These foundations arise from what we know about the effective use of educational technology to improve opportunity, equity, and outcomes for students and also relate to the new Blueprint.

Foundation 1: Center people (parents, educators, and students)

Education-focused AI policies at the federal, state, and district levels will be needed to guide and empower local and individual decisions about which technologies to adopt and use in schools and classrooms. Consider what is happening in everyday lives. Many of us use AI-enabled products because they are often better and more convenient. For example, few people want to use paper maps anymore; people find that technology helps us plan the best route to a destination more efficiently and conveniently. And yet, people often do not realize how much privacy they are giving up when they accept AI-enabled systems into their lives. AI will bring privacy and other risks that are hard to address only via individual decision making; additional There should be clear limits on the ability to collect, use, transfer, and maintain our personal data, including limits on targeted advertising. These limits should put the burden on platforms to minimize how much information they collect, rather than burdening Americans with reading fine print.⁸ As protections are developed, we recommend that policies center people, not machines. To this end, a first

recommendation in this document (in the next section) is an emphasis on AI with humans in the loop. Teachers, learners, and others need to retain their agency to decide what patterns mean and to choose courses of action. The idea of humans in the loop builds on the concept of “Human Alternatives, Consideration, and Fallback” in the Blueprint and ethical concepts used more broadly in evaluating AI, such as preserving human dignity. A top policy priority must be establishing human in the loop as a requirement in educational applications, despite contrary pressures to use AI as an alternative to human decision making. Policies should not hinder innovation and improvement, nor should they be burdensome to implement. Society needs an education-focused AI policy that protects civil rights and promotes democratic values in the building, deployment, and governance of automated systems to be used across the many decentralized levels of the American educational system.

Foundation 2: Advance Equity

“AI brings educational technology to an inflection point. We can either increase disparities or shrink them, depending on what we do now.” —Dr. Russell Shilling A recent Executive Order⁹ issued by President Biden sought to strengthen the connection among racial equity, education and AI, stating that “members of underserved communities—many of whom have endured generations of discrimination and disinvestment—still confront significant barriers to realizing the full promise of our great Nation, and the Federal Government has a responsibility to remove these barriers” and that the Federal Government shall both “pursue educational equity so that our Nation’s schools put every student on a path to success” and also “root out bias in the design and use of new technologies, such as artificial intelligence.” A specific vision of equity, such as described in the Department’s recent report, Advancing Digital Equity for All¹⁰ is essential to policy discussion about AI in education. This report defines digital equity as “the condition in which individuals and communities have the information technology capacity that is needed for full participation in the society and economy of the United States.” Issues related to racial equity and unfair bias were at the heart of every listening session we held. In particular, we heard a conversation that was increasingly attuned to issues of data quality and the consequences of using poor or inappropriate data in AI systems for education. Datasets are used to develop AI, and when they are non-representative or contain undesired associations or patterns, resulting AI models may act unfairly in how they detect patterns or automate decisions. Systematic, unwanted unfairness in how a computer detects patterns or automates decisions is called “algorithmic bias.” Algorithmic bias

could diminish equity at scale with unintended discrimination. As this document discussed in the Formative Assessment section, this is not a new conversation. For decades, constituents have rightly probed whether assessments are unbiased and fair. Just as with assessments, whether an AI model exhibits algorithmic bias or is judged to be fair and trustworthy is critical as local school leaders make adoption decisions about using AI to achieve their equity goals. We highlight the concept of “algorithmic discrimination” in the Blueprint. Bias is intrinsic to how AI algorithms are developed using historical data, and it can be difficult to anticipate all impacts of biased data and algorithms during system design. The Department holds that biases in AI algorithms must be addressed when they introduce or sustain unjust discriminatory practices in education. For example, in postsecondary education, algorithms that make enrollment decisions, identify students for early intervention, or flag possible student cheating on exams must be interrogated for evidence of unfair discriminatory bias—and not only when systems are designed, but also later, as systems become widely used.

Foundation 3: Ensure Safety, Ethics, and Effectiveness

A central safety argument in the Department’s policies is the need for data privacy and security in the systems used by teachers, students, and others in educational institutions. The development and deployment of AI requires access to detailed data. This data goes beyond conventional student records (roster and gradebook information) to detailed information about what students do as they learn with technology and what teachers do as they use technology to teach. AI’s dependence on data requires renewed and strengthened attention to data privacy, security, and governance (as also indicated in the Blueprint). As AI models are not generally developed in consideration of educational usage or student privacy, the educational application of these models may not be aligned with the educational institution’s efforts to comply with federal student privacy laws, such as FERPA, or state privacy laws. Further, educational leaders are committed to basing their decisions about the adoption of educational technology on evidence of effectiveness—a central foundation of the Department’s policy. For example, the requirement to base decisions on evidence also arises in the Elementary and Secondary Education Act (ESEA), as amended, which introduced four tiers of evidence (see Figure 2). Our nation’s research agencies, including the Institute of Education Sciences, are essential to producing the needed evidence. The Blueprint calls for evidence of effectiveness, but the education sector is ahead of that game: we need to insist that AI-enhanced edtech rises to meet ESEA standards as well.

Foundation 4: Promote Transparency

The central role of complex AI models in a technology’s detection of patterns and implementation of automation is an important way in which AI-enabled applications, products, and services will be different from conventional edtech. The Blueprint introduces the need for transparency about AI models in terms of disclosure (“notice”) and explanation. In education, decision makers will need more than notice—they will need to understand how AI models work in a range of general educational use cases, so they can better anticipate limitations, problems, and risks. AI models in edtech will be approximations of reality and, thus, constituents can always ask these questions: How precise are the AI models? Do they accurately capture what is most important? How well do the recommendations made by an AI model fit educational goals? What are the broader implications of using AI models at scale in educational processes? Building on what was heard from constituents, the sections of this report develop the theme of evaluating the quality of AI systems and tools using multiple dimensions as follows:

- About AI: AI systems and tools must respect data privacy and security. Humans must be in the loop.
- Learning: AI systems and tools must align to our collective vision for high-quality learning, including equity.
- Teaching: AI systems and tools must be inspectable, explainable, and provide human alternatives to AI-based suggestions; educators will need support to exercise professional Formative Assessment: AI systems and tools must minimize bias, promote fairness, and avoid additional testing time and burden for students and teachers.
- Research and Development: AI systems and tools must account for the context of teaching and learning and must work well in educational practice, given variability in students, teachers, and settings.
- Recommendations: Use of AI systems and tools must be safe and effective for students. They must include algorithmic discrimination protections, protect data privacy, provide notice and explanation, and provide a recourse to humans when problems arise. The people most affected by the use of AI in education must be part of the development of the AI model, system, or tool, even if this slows the pace of adoption. We return to the idea that these considerations fit together in a comprehensive perspective on the quality of AI models in the Recommendations section.

Overview of Document

We begin in the next section by elaborating a definition of AI, followed by addressing learning, teaching, assessment, and research and development. Organizing key insights by these topics keeps us focused on exploring implications for improving educational opportunity and outcomes for

students throughout the report. Within these topics, three important themes are explored: 1. Opportunities and Risks. Policies should focus on the most valuable educational advances while mitigating risks. 2. Trust and Trustworthiness. Trust and safeguarding are particularly important in education because we have an obligation to keep students out of harm's way and safeguard their learning experiences. 3. Quality of AI Models. The process of developing and then applying a model is at the heart of any AI system. Policies need to support evaluation of the qualities of AI models and their alignment to goals for teaching and learning during the processes of educational adoption and use.

What is AI?

Our preliminary definition of AI as automation based on associations requires elaboration. Below we address three additional perspectives on what constitutes AI. Educators will find these different perspectives arise in the marketing of AI functionality and are important to understand when evaluating edtech systems that incorporate AI. One useful glossary of AI for Education terms is the CIRCLS Glossary of Artificial Intelligence Terms for Educators. 11 AI is not one thing but an umbrella term for a growing set of modeling capabilities, as visualized in Figure 3.

Perspective: Human-Like Reasoning

“The theory and development of computer systems able to perform tasks normally requiring human intelligence such as, visual perception, speech recognition, learning, decision-making, and natural language processing.” 13 Broad cultural awareness of AI may be traced to the landmark 1968 film “2001: A Space Odyssey”—in which the “Heuristically-programmed ALgorithmic” computer, or “HAL,” converses with astronaut Frank. HAL helps Frank pilot the journey through space, a job that Frank could not do on his own. However, Frank eventually goes outside the spacecraft, HAL takes over control, and this does not end well for Frank. HAL exhibits human-like behaviors, such as reasoning, talking, and acting. Like all applications of AI, HAL can help humans but also introduces unanticipated risks—especially since AI reasons in different ways and with different limitations than people do. The idea of “human-like” is helpful because it can be a shorthand for the idea that computers now have capabilities that are very different from the capabilities of early edtech applications. Educational applications will be able to converse with students and teachers, co-pilot how activities unfold in classrooms, and take actions that impact

students and teachers more broadly. There will be both opportunities to do things much better than we do today and risks that must be anticipated and addressed. The “human-like” shorthand is not always useful, however, because AI processes information differently from how people process information. When we gloss over the differences between people and computers, we may frame policies for AI in education that miss the mark.

Perspective: An Algorithm that Pursues a Goal

“Any computational method that is made to act independently towards a goal based on inferences from theory or patterns in data.”¹⁴ This second definition emphasizes that AI systems and tools identify patterns and choose actions to achieve a given goal. These pattern recognition capabilities and automated recommendations will be used in ways that impact the educational process, including student learning and teacher instructional decision making. For example, today’s personalized learning systems may recognize signs that a student is struggling and may recommend an alternative instructional sequence. The scope of pattern recognition and automated recommendations will expand Correspondingly, humans must determine the types and degree of responsibility we will grant to technology within educational processes, which is not a new dilemma. For decades, the lines between the role of teachers and computers have been discussed in education, for example, in debates using terms such as “computer-aided instruction,” “blended instruction,” and “personalized learning.” Yet, how are instructional choices made in systems that include both humans and algorithms? Today, AI systems and tools are already enabling the adaptation of instructional sequences to student needs to give students feedback and hints, for example, during mathematics problem solving or foreign language learning. This discussion about the use of AI in classroom pedagogy and student learning will be renewed and intensify as AI-enabled systems and tools advance in capability and become more ubiquitous. Let’s start with another simple example. When a teacher says, “Display a map of ancient Greece on the classroom screen,” an AI system may choose among hundreds of maps by noting the lesson objectives, what has worked well in similar classrooms, or which maps have desirable features for student learning. In this case, when an AI system suggests an instructional resource or provides a choice among a few options, the instructor may save time and may focus on more important goals. However, there are also forms of AI-enabled automation that the classroom instructor may reject, for example, enabling an AI system or tool to select the most appropriate and relevant readings for students associated with a historical event. In this case, an educator may choose not to utilize AI-enabled

systems or tools given the risk of AI creating false facts (“hallucinating”) or steering students toward inaccurate depictions of historical events found on the internet. Educators will be weighing benefits and risks like these daily. Computers process theory and data differently than humans. AI’s success depends on associations or relationships found in the data provided to an algorithm during the AI model development process. Although some associations may be useful, others may be biased or inappropriate. Finding bad associations in data is a major risk, possibly leading to algorithmic discrimination. Every guardian is familiar with the problem: A person or computer may say, “Our data suggests your student should be placed in this class,” and the guardian may well argue, “No, you are using the wrong data. I know my child better, and they should instead be placed in another class.” This problem is not limited exclusively to AI systems and tools, but the use of AI models can amplify the problem when a computer uses data to make a recommendation because it may appear to be more objective and authoritative, even if it is not. Although this perspective can be useful, it can be misleading. A human view of agency, pursuing goals, and reasoning includes our human abilities to make sense of multiple contexts. For example, a teacher may see three students each make the same mathematical error but recognize that one student has an Individualized Education Program to address vision issues, another misunderstands a mathematical concept, and a third just experienced a frustrating interaction on the playground; the same instructional decision is therefore not appropriate. However, AI systems often lack data and judgement to appropriately include context as they detect patterns and automate decisions. Further, case studies show that technology has the potential to quickly derail from safe to unsafe or from effective to ineffective when the context shifts even slightly. For this and other reasons, people must be involved in goal setting, pattern analysis, and decision-making.¹⁵

Perspective: Intelligence Augmentation

Foundation #1 (above) keeps humans in the loop and positions AI systems and tools to support human reasoning. “Intelligence Augmentation” (IA)¹⁷ centers “intelligence” and “decision making” in humans but recognizes that people sometimes are overburdened and benefit from assistive tools. AI may help teachers make better decisions because computers notice patterns that teachers can miss. For example, when a teacher and student agree that the student needs reminders, an AI system may provide reminders in whatever form a student likes without adding to the teacher’s workload. Intelligence Automation (IA) uses the same basic capabilities of AI, employing associations in data to notice patterns, and, through automation, takes actions based on

those patterns. However, IA squarely focuses on helping people in human activities of teaching and learning, whereas AI tends to focus attention on what computers can do.

Definition of “Model”

The above perspectives open a door to making sense of AI. Yet, to assess AI meaningfully, constituents must consider specific models and how they are developed. In everyday usage, the term “model” has multiple definitions. We clarify our intended meaning, which is a meaning similar to “mathematical model,” below. (Conversely, note that “model” as used in “AI model” is unlike the usage in “model school” or “instructional model” as AI model is not a singular case created by experts to serve as an exemplar.) AI models are like financial models: an approximation of reality that is useful for identifying patterns, making predictions, or analyzing alternative decisions. In a typical middle school math curriculum, students use a mathematical model to analyze which of two cell phone plans is better. Financial planners use this type of model to provide guidance on a retirement portfolio. At its heart, AI is a highly advanced mathematical toolkit for building and using models. Indeed, in well-known chatbots, complex essays are written one word at a time. The underlying AI model predicts which next words would likely follow the text written so far; AI chatbots use a very large statistical model to add one likely word at a time, thereby writing surprisingly coherent essays. When we ask about the model at the heart of AI, we begin to get answers about “what aspects of reality does the model approximate well?” and “how appropriate is it to the decision to be made?” One could similarly ask about algorithms—the specific decision-making processes that an AI model uses to go from inputs to outputs. One could also ask about the quality of the data used to build the model—for example, how representative is that data? Switching among three terms—models, algorithms, and data—will become confusing. Because the terms are closely related, we’ve chosen to focus on the concept of AI models. We want to bring to the fore the idea that every AI model is incomplete, and it’s important to know how well the AI model fits the reality we care about, where the model will break down, and how. Sometimes people avoid talking about the specifics of models to create a mystique. Talking as though AI is unbounded in its potential capabilities and a nearly perfect approximation to reality can convey an excitement about the possibilities of the future. The future, however, can be oversold. Similarly, sometimes people stop calling a model AI when its use becomes commonplace, yet such systems are still AI models with all of the risks discussed here. We need to know exactly when and where AI models fail to align to visions for teaching and learning.

Insight: AI Systems Enable New Forms of Interaction

AI models allow computational processes to make recommendations or plans and also enable them to support forms of interaction that are more natural, such as speaking to an assistant. AI-enabled educational systems will be desirable in part due to their ability to support more natural interactions during teaching and learning. In classic edtech platforms, the ways in which teachers and students interact with edtech are limited. Teachers and students may choose items from a menu or in a multiple-choice question. They may type short answers. They may drag objects on the screen or use touch gestures. The computer provides outputs to students and teachers through text, graphics, and multimedia. Although these forms of inputs and outputs are versatile, no one would mistake this style of interaction with the way two people interact with one another; it is specific to human-computer interaction. With AI, interactions with computers are likely to become more like human-to-human interactions (see Figure 4). A teacher may speak to an AI assistant, and it may speak back. A student may make a drawing, and the computer may highlight a portion of the drawing. A teacher or student may start to write something, and the computer may finish their sentence—as when today’s email programs can complete thoughts faster than we can type them. Additionally, the possibilities for automated actions that can be executed by AI tools are expanding. Current personalization tools may automatically adjust the sequence, pace, hints, or trajectory through learning experiences.¹⁸ Actions in the future might look like an AI system or tool that helps a student with homework¹⁹ or a teaching assistant that reduces a teacher’s workload by recommending lesson plans that fit a teacher’s needs and are similar to lesson plans a teacher previously liked.²⁰ Further, an AI-enabled assistant may appear as an additional “partner” in a small group of students who are working together on a collaborative assignment.²¹ An AI-enabled tool may also help teachers with complex classroom routines.²² For example, a tool may help teachers with orchestrating²³ the movement of students from a full class discussion into small groups and making sure each group has the materials needed to start their work.

Key Recommendation: Human in the Loop AI

Many have experienced a moment where technology surprised them with an uncanny ability to recommend what feels like a precisely personalized product, song, or even phrase to complete a sentence in a word processor such as the one being used to draft this document. Throughout this supplement, we talk about specific, focused applications where AI systems may bring value (or risks) into education. At no point do we intend to imply that AI can replace a teacher, a guardian,

or an educational leader as the custodian of their students' learning. We talk about the limitations of models in AI and the conversations that educational constituents need to have about what qualities they want AI models to have and how they should be used. These limitations lead to our first recommendation: that we pursue a vision of AI where humans are in the loop. That means that people are part of the process of noticing patterns in an educational system and assigning meaning to those patterns. It also means that teachers remain at the helm of major instructional decisions. It means that formative assessments involve teacher input and decision making, too. One loop is the cycle of recognizing patterns in what students do and selecting next steps or resources that could support their learning. Other loops involve teachers planning and reflecting on lessons. Response to Intervention is another well-known type of loop. The idea of humans in the loop is part of our broader discussions happening about AI and society, not just AI in education. Interested readers could look for more on human-centered AI, responsible AI, value-sensitive AI, AI for social good, and other similar terms that ally with humans in the loop, such as "human-centered AI." Exercising judgement and control in the use of AI systems and tools is an essential part of providing the best opportunity to learn for all students—especially when educational decisions carry consequence. AI does not have the broad qualities of contextual judgment that people do. Therefore, people must remain responsible for the health and safety of our children, for all students' educational success and preparation for their futures, and for creating a more equitable and just society.

Learning

The Department's long-standing edtech vision sees students as active learners; students participate in discussions that advance their understanding, use visualizations and simulations to explain concepts as they relate to the real world, and leverage helpful scaffolding and timely feedback as they learn. Constituents want technology to align to and build on these and other research-based understandings of how people learn. Educators can draw upon two books titled *How People Learn* and *How People Learn II* by the National Academies of Sciences, Engineering, and Medicine for a broad synthesis of what we know about learning.²⁴ As we shape AI-enhanced edtech around research-based principles, a key goal must be to strengthen and support learning for those who have experienced unfavorable circumstances for learning, such as caused by the COVID-19

pandemic or by broader inequities. And we must keep a firm eye toward the forms of learning that will most benefit learners in their future lives in communities and workplaces. Examples of AI supporting learning principles in this section include the following: AI-based tutoring for students as they solve math problems (based on cognitive learning theories), adapting to learners with special needs (based on the Universal Design for Learning framework and related theories), and AI support for effective student teamwork (based on theories in the field called “Computer Supported Collaborative Learning”).

Insight: AI Enables Adaptivity in Learning

Adaptivity has been recognized as a key way in which technology can improve learning.²⁵ AI can be a toolset for improving the adaptivity of edtech. AI may improve a technology’s ability to meet students where they are, build on their strengths, and grow their knowledge and skills. Because of AI’s powers of work with natural forms of input and the foundational strengths of AI models (as discussed in the What is AI? section), AI can be an especially strong toolkit for expanding the adaptivity provided to students. And yet, especially with AI, adaptivity is always more specific and limited than what a broad phrase like “meet students where they are” might suggest. Core limits arise from the nature of the model at the heart of any specific AI-enabled system. Models are approximations of reality. When important parts of human learning are left out of the model or less fully developed, the resulting adaptivity will also be limited, and the resulting supports for learning may be brittle or narrow. Consequently, this section on Learning focuses on one key concept: Work toward AI models that fit the fullness of visions for learning—and avoid limiting learning to what AI can currently model well. AI models are demonstrating greater skills because of advances in what are called “large language models” or sometimes “foundational models.” These very general models still have limits. For example, generative AI models discussed in the mainstream news can quickly generate convincing essays about a wide variety of topics while other models can draw credible images based on just a few prompts. Despite the excitement about foundational models, experts in our listening sessions warned that AI models are narrower than visions for human learning and that designing learning environments with these limits in mind remains very important. The models are also brittle and can’t perform well when contexts change. In addition, they don’t have the same “common sense” judgment that people have, often responding in ways that are unnatural or incorrect.²⁶ Given the unexpected ways in which foundational models miss the mark, keeping humans in the loop remains highly important.

Intelligent Tutoring Systems: An Example of AI Models

One long-standing type of AI-enabled technology is an Intelligent Tutoring System (ITS).²⁷ In an early success, scientists were able to build accurate models of how human experts solve mathematical problems. The resulting model was incorporated into a system that would observe student problem solving as they worked on mathematical problems on a computer. Researchers who studied human tutors found that feedback on specific steps (and not just right or wrong solutions) is a likely key to why tutoring is so effective.²⁸ For example, when a student diverged from the expert model, the system gave feedback to help the student get back on track.²⁹ Importantly, this feedback went beyond right or wrong, and instead, the model was able to provide feedback on specific steps of a solution process. A significant advancement of AI, therefore, can be its ability to provide adaptivity at the step-by-step level and its ability to do so at scale with modest cost. As a research and development (R&D) field emerged to advance ITS, the work has gone beyond mathematics problems to additional important issues beyond step-by-step problem solving. In the early work, some limitations can be observed. The kinds of problems that an ITS could support were logical or mathematical, and they were closed tasks, with clear expectations for what a solution and solution process should look like. Also, the “approximation of reality” in early AI models related to cognition and not to other elements of human learning, for example, social or motivational aspects. Over time, these early limitations have been addressed in two ways: by expanding the AI models and by involving humans in the loop, a perspective that is also important now. Today, for example, if an ITS specializes in feedback as a student practices, a human teacher could still be responsible for motivating student engagement and self-regulation along with other aspects of instruction. In other contemporary examples, the computer ITS might focus on problem solving practice, while teachers work with students in small groups. Further, students can be in the loop with AI, as is the case with “open learner models”—a type of AI-enabled system that provides information to support student self-monitoring and reflection. Although R&D along the lines of an ITS should not limit the view of what’s possible, such an example is useful because so much research and evaluation has been done on the ITS approach. Researchers have looked across all the available high-quality studies in a meta-analysis and concluded that ITS approaches are effective.³¹ Right now, many school systems are looking at high-intensity human tutoring to help students with unfinished learning. Human tutoring is very expensive, and it is hard to find enough high-quality human tutors. With regard to large-scale

needs, if it is possible for an ITS to supplement what human tutors do, it might be possible to extend beyond the amount of tutoring that people can provide to students.

Important Directions for Expanding AI-Based Adaptivity

Adaptivity is sometimes referred to as “personalization.” Although this is a convenient term, many observers have noted how imprecise it is.³² For some educators, personalization means giving learners “voice and choice,” and for others it means that a learning management system recommends an individual “playlist” of activities to each student. Hidden in that imprecision is the reality that many edtech products that personalize do so in limited ways. Adjusting the difficulty and the order of lesson materials are among the two most common ways that edtech products adapt. And yet, any teacher knows there is more to supporting learning than adjusting the difficulty and sequence of materials. For example, a good teacher can find ways to engage a student by connecting to their own past experiences and can shape explanations until they really connect in an “aha!” moment for that student. When we say, “meet the learner where they are,” human teachers bring a much more complete picture of each learner than most available edtech. The teacher is also not likely to “over personalize” (by performing like an algorithm that only presents material for which the learner has expressed interest), thereby limiting the student’s exposure to new topics. The nature of “teachable moments” that a human teacher can grasp is broader than the teachable moments today’s AI models grasp. In our listening sessions, we heard many ways in which the core models in an AI system must be expanded. We discuss these below.

1. From deficit-based to asset-oriented. Listening session attendees noted that the rhetoric around adaptivity has often been deficit-based; technology tries to pinpoint what a student is lacking and then provides instruction to fill that specific gap. Teachers also orient to students’ strengths; they find competencies or “assets” a student has and use those to build up the students’ knowledge. AI models cannot be fully equitable while failing to recognize or build upon each student’s sources of competency. AI models that are more asset-oriented would be an advance.
2. From individual cognition to including social and other aspects of learning. The existing adaptivity rhetoric has also tended to focus on individualized learning and mostly on cognitive elements of learning, with motivational and other elements only brought in to support the cognitive learning goals. Attendees observe that their vision for learning is broader than cognition. Social learning is important, for example, especially for students to learn to reason, explain, and justify. For students who are learning English, customized and adaptive support for improving language skills while learning

curricular content is clearly important. Developing self-regulation skills is also important. A modern vision of learning is not individualistic; it recognizes that students learn in groups and communities too.

3. From neurotypical to neurodiverse learners. AI models could help in including neurodiverse learners (students who access, process, and interact with the world in less common ways than “neurotypical” students) who could benefit from different learning paths and from forms of display and input that fit their strengths. Constituents want AI models that can support learning for neurodiverse learners and learners with disabilities. Thus, they want AI models that can work with multiple paths to learning and multiple modalities of interaction. Such models should be tested for efficacy, to guard against the possibility that some students could be assigned a “personalized” but inadequate learning resource. In addition, some systems for neurodiverse students are presently underutilized, so designs that support intended use will also be important.

4. From fixed tasks to active, open, and creative tasks. As mentioned above, AI models are historically better at closed tasks like solving a math problem or logical tasks like playing a game. In terms of life-wide and lifelong opportunities, we value learning how to succeed at open-ended and creative tasks that require extended engagement from the learner, and these are often not purely mathematical or logical. We want students to learn to invent and create innovative approaches. We want AI models that enable progress on open, creative tasks.

5. From correct answers to additional goals. At the heart of many adaptivity approaches now on the market, the model inside the technology counts students' wrong answers and decides whether to speed up, slow down, or offer a different type of learning support. Yet, right and wrong answers are not the only learning goals. We want students to learn how to self-regulate when they experience difficulties in learning, for example, such as being able to persist in working on a difficult problem or knowing how and when to ask for help. We want learners to become skilled in teamwork and in leading teams. As students grow, we want them to develop more agency and to be able to act on their own to advance toward their own learning goals. Listing every dimension of expansion that we heard in our listening sessions is beyond the scope of this report. Some additional dimensions are presented in the following sections on Teaching, Assessment, and Research. For example, in Research, we discuss all the ways in which AI systems have trouble with context—context that humans readily grasp and consider. Overall, constituents in the listening sessions realized we need an ambitious outlook on learning to respond to the future today's learners face. Constituents were concerned about ways in which AI might narrow learning. For example, if the incorporation of AI into education slowed

attention to students' skills on creative, open-ended tasks and their ability to lead and collaborate in teams, then school districts may be less able to realize their students' progress in relation to a Portrait of a Graduate who excels in communication and other skills valued in communities and careers. Constituents reminded us that as we conceptualize what we want AI in edtech to accomplish, we must start and constantly revisit a human-centered vision of learning.

A Duality: Learning With and About AI

As AI is brought into schools, two broad perspectives about AI in education arise: (1) AI in support of student learning; and (2) support for learning about AI and related technologies. So far, we've discussed AI systems and tools to support student learning and mastery of subjects like mathematics and writing. Yet, it is also important that students learn about AI, critically examine its presence in education and society, and determine its role and value in their own lives and careers. We discuss risks across each section in this report. Here, it is important for students to become more aware of and savvy to the risks of AI—including risks of bias and surveillance—as they appear in all elements of their lives. In the recent past, schools have supported students' understanding of cybersecurity, for example. AI will bring new risks, and students need to learn about them. We are encouraged by efforts we've seen underway that would give students opportunities to learn about how AI works while also giving them opportunities to discuss relevant topics like privacy and security.³³ Other learning goals are noted in the K-12 Computer Science Framework. We've seen that students can begin learning about AI in elementary, middle, and high school. They can use AI to design simulations and products that they find exciting. And we've seen that students want to talk about the ethics of products they experience in their everyday lives and have much to say about the kinds of products they'd like to see or not see in school. (And later, in the Research section, we note the desire for co-design processes that involve students in creating the next generation of AI-enabled edtech). Overall, it's important to balance attention to using AI to support learning and giving students opportunities to learn about AI.

A Challenge: Systems Thinking About AI in Education

As AI expands into the educational system, our listening session attendees reminded us that it will be entering parts or locations of the system that are presently dysfunctional. AI is certainly not a fix for broken systems, and instead, must be used with even more care when the systems' context is unstable or uncertain. As discussed previously, because AI systems and tools do not fully align with goals for learning, we have to design educational settings to situate AI in the right place,

where educators and other adults can make effective use of these tools for teaching and learning. Within the ITS example, we saw that AI could make learning by practicing math problems more effective, and a whole curricular approach might include roles for teachers that emphasize mathematical practices like argumentation and modeling. Further, small-group work is likely to remain important: Students might work in small groups to use mathematics to predict or justify as they work on responding to a realistic challenge. At the present, one “right place” for people, and not AI, is understanding how learning can be culturally responsive and culturally sustaining, as AI is not even close to being ready to connect learning to the unique strengths in a student’s community and family

Open Questions About AI for Learning

With advances occurring in the foundations for AI, opportunities to use AI in support of learning are rapidly expanding. As we explore these opportunities, the open questions below deserve ongoing attention:

- To what extent is AI enabling adaptation to students’ strengths and not just deficits? Is AI enabling improved support for learners with disabilities and English language learners?
- How are youth voices involved in choosing and using AI for learning?
- Is AI leading to narrower student activities (e.g., procedural math problems), or the fuller range of activities highlighted in the National Educational Technology Plan (NETP), which emphasizes features such as personalized learning, project-based learning, learning from visualizations, simulations, and virtual reality, as well as learning across school, community, and familial settings?
- Is AI supporting the whole learner, including social dimensions of learning such as enabling students to be active participants in small group and collaborative learning? For example, does AI contribute to aspects of student collaboration we value like shared attention, mutual engagement, peer help, self-regulation, and building on each other’s contributions?
- When AI is used, are students’ privacy and data protected? Are students and their guardians informed about what happens with their data?
- How strong are the processes or systems for monitoring student use of AI for barriers, bias, or other undesirable consequences of AI use by learners? How are emergent issues addressed?
- Is high-quality research or evaluations about the impacts of using the AI system for student learning available? Do we know not only whether the system works but for whom and under what conditions?

Key Recommendation: Seek AI Models Aligned to a Vision for Learning

We've called attention to how advances in AI are important to adaptivity but also to ways in which adaptivity is limited by the model's inherent quality. We noted that a prior wave of edtech used the term "personalized" in differing ways, and it was often important to clarify what personalization meant for a particular product or service. Thus, our key recommendation is to tease out the strengths and limitations of AI models inside forthcoming edtech products and to focus on AI models that align closely to desired visions of learning. AI is now advancing rapidly, and we should differentiate between products that have simple AI-like features inside and products that have more sophisticated AI models. Looking at what's happening in research and development, we can see significant effort and push toward overcoming these limitations. We noted that decision makers need to be careful about selecting AI models that might narrow their vision for learning, as general artificial intelligence does not exist. And because AI models will always be narrower than real world experience, we need to proceed with systems thinking in which humans are in the loop, with the strengths and weaknesses of the specific educational system considered. We hold that the full system for learning is broader than its AI component.

Teaching

Teachers have long envisioned many things that technology could make possible for teachers, their classrooms, and their students but not the changes wrought by the recent pandemic. Today, nearly all teachers have experienced uses of technologies for instruction that no one anticipated. Some of those experiences were positive, and others were not. All of the experiences provide an important context as we think further about teaching and technology. There is a critical need to focus on addressing the challenges teachers experience. It must become easier for teachers to do the amazing work they always do. We must also remember why people choose the teaching profession and ensure they can do the work that matters. This section discusses examples of AI supporting teachers and teaching including these concepts: AI assistants to reduce routine teaching burdens; AI that provides teachers with recommendations for their students' needs and extends their work with students; and AI that helps teachers to reflect, plan, and improve their practice.

Always Center Educators in Instructional Loops

To succeed with AI as an enhancement to learning and teaching, we need to always center educators (ACE). Practically speaking, practicing “ACE in AI” means keeping a humanistic view of teaching front and center. ACE leads the Department to confidently respond “no” when asked “will AI replace teachers?” ACE is not just about making teachers’ jobs easier but also making it possible to do what most teachers want to do. That includes, for example, understanding their students more deeply and having more time to respond in creative ways to teachable moments. To bring more precision to how and where we should center educators, we return to our advocacy for human in the loop AI and ask, what are the loops in which teachers should be centered? Figure 5 suggests three key loops (inspired by research on adaptivity loops³⁴): 1. The loop in which teachers make moment-to-moment decisions as they do the immediate work of teaching. 2. The loop in which teachers prepare for, plan, and reflect on teaching, which includes professional development. 3. The loop in which teachers participate in decisions about the design of AI-enabled technologies, participate in selecting the technologies, and shape the evaluation of technologies—thus setting a context for not only their own classroom but those of fellow teachers as well. Please note that in the next section, on Formative Assessment, we also discuss teachers’ important role in feedback loops that support students and enable school improvement. That section also includes a discussion of the concepts of “bias” and “fairness,” which are important to teachers.

Insight: Using AI to Improve Teaching Jobs

The job of teaching is notoriously complex, with teachers making thousands of decisions each day. Teachers participate in classroom processes, in interactions with students beyond classrooms, in work with fellow teachers, and in administrative functions. They also are part of their communities and thus are expected to interact with families and caregivers alerts and notifications about events. Selecting music that we want to hear used to be a multistep process (even with digital music), and now we can speak the name of a song we want to hear, and it plays. Likewise, mapping a journey used to require a cumbersome study of maps, but now cell phones let us choose among several transportation options to reach a destination. Why can’t teachers be supported to notice changing student needs and provided with supports to enact a technology-rich lesson plan? Why can’t they more easily plan their students’ learning journeys? When things change in a classroom, as they always do, why don’t the tools of the classroom make it easier for teachers to adapt to student strengths and needs on the fly? A report by McKinsey³⁶ first suggested that AI’s initial benefit

could be to improve teaching jobs by reducing low-level burdens in administrative or clerical work (Figure 6). The report also suggests that recovered time from AI-enabled technology should be rededicated toward more effective instruction—particularly, outcomes such as reducing the average 11 hours of weekly preparation down to only six. We highlight these opportunities and two others below.

1. Handling low-level details to ease teaching burdens and increase focus on students. A good teacher must master all levels of details, big and small. When working with a particular student, the teacher may wish to later send that student a helpful learning resource. How will they remember to send it? A voice assistant or other forms of an AI assistant could make it easier to stay organized by categorizing simple voice notes for teachers to follow up on after a classroom session ends. We are beginning to see AI-enabled voice assistants in the market, and they could do many simple tasks so that the teachers can stay focused on students. These tasks can include record-keeping, starting and stopping activities, controlling displays, speakers, and other technologies in the classroom, and providing reminders. Many workers may eventually use assistants to make their jobs easier, and teachers are the most deserving of efforts to ease their jobs now.
2. Extending beyond the teacher's availability with their students but continuing to deliver on the teacher's intent. Teachers almost always want to do more with each student than they can, given the limited number of hours before the next school day. A teacher may wish to sit with the student as they practice 10 more math problems, giving them ongoing support and feedback. If the teacher can sit with the student for only three problems, perhaps they could delegate to an AI-enabled learning system to help with the rest. Teachers cannot be at their best if on call at all hours to help with homework, but perhaps they can indicate what types of supports, hints, and feedback they want students to receive while studying after school hours. An AI assistant can ensure that students have that support wherever and whenever they do homework or practice skills on their own. Teachers may wish to provide more extensive personal notes to families/caregivers, and perhaps an AI assistant could help with drafts based on students' recent classroom work. Then, the teacher could review the AI-generated comments and quickly edit where needed before returning it to the student for another draft. AI tools might also help teachers with language translation so they can work with all parents and caregivers of their students. AI tools might also help teachers with awareness. For example, in the next section, Formative Assessment, we note that teachers can't always know what's going on for each student and in each small group of students; emerging products might signal to the teacher when a student or teacher may need some more personal

attention.

3. Making teacher professional development more productive and fruitful. Emerging products already enable a teacher to record her classroom and allow an AI algorithm to suggest highlights of the classroom discussion worth reviewing with a professional development coach.³⁷ AI can compute metrics, such as whether students have been talking more or less, which are difficult for a teacher to calculate during a lesson.³⁸ For teachers who want to increase student engagement, these metrics can be a valuable tool. Classroom simulation tools are also emerging and can enable teachers to practice their skills in realistic situations.³⁹ Simulators can include examples of teaching from a real classroom while changing the faces and voices of the participants so that teaching situations can be shared and discussed among teachers without revealing identities. Note the emphasis above on what listening-session panelist Sarah Hampton said about the human touch. Teachers will feel that AI is helping them teach with a focus on their human connection to their students when the necessary (but less meaningful) burdens of teaching are lessened. In Figure 7, below, see concerns that teachers raised about AI during listening sessions.

Preparing and Supporting Teachers in Planning and Reflecting

ACE also means preparing teachers to take advantage of possibilities like those listed above and more. In the Research section, we highlight how pre-service education still tends to compartmentalize and inadequately address the topic of technology. That section suggests a need to invest in research about how to deeply integrate technology in pre-service teacher training programs. In-service teachers, too, will need professional development to take advantage of opportunities that AI can provide, like those presented in the Teaching section. Professional development will need to be balanced not only to discuss opportunities but also to inform teachers of new risks, while providing them with tools to avoid the pitfalls of AI. By nature, teaching requires significant time in planning as well to account for the breadth of needs across their rosters—especially for inclusive learning environments and students with IEPs and 504 plans. AI could help teachers with recommendations that are tuned to their situation and their ways of practicing teaching and support with adapting found materials to fit their exact classroom needs. For students with an IEP, AI could help with finding components to add to lesson plans to fully address standards and expectations and to meet each student's unique requirements. Even beyond finding components, AI might help adapt standardized resources to better fit specific needs—for example, providing a voice assistant that allows a student with a visual difficulty to hear material and respond to it or permitting a group of students to present their project using American Sign

Language (ASL) which could be audibly voiced for other students using an AI ASL-to-Spoken-English translation capability. Indeed, coordinating IEPs is time-consuming work that might benefit from supportive automation and customized interactivity that can be provided by AI. Reflection is important too. In the bustle of a classroom, it is sometimes difficult to fully understand what a student is expressing or what situations lead to certain positive or negative behaviors. Again, context is paramount. In the moment, teachers may not be aware of external events that could shape their understanding of how students are showing up in their classrooms. Tools that notice patterns and suggest ways to share information might help students and teachers communicate more fully about strengths and needs.

Designing, Selecting, and Evaluating AI Tools

The broadest loop teachers should be part of is the loop that determines what classroom tools do and which tools are available. Today, teachers already play a role in designing and selecting technologies. Teachers can weigh in on usability and feasibility. Teachers examine evidence of efficacy and share their findings with other school leaders. Teachers already share insights on what is needed to implement technology well. While these concerns will continue, AI will raise new concerns too. For example, the following Formative Assessment section raises concerns about bias and fairness that can lead to algorithmic discrimination. Those concerns go beyond data privacy and security; they raise attention to how technologies may unfairly direct or limit some students' opportunities to learn. A key takeaway here is that teachers will need time and support so they can stay abreast of both the well-known and the newer issues that are arising and so they can fully participate in design, selection, and evaluation processes that mitigate risks.

Challenge: Balancing Human and Computer Decision-Making

One major new challenge with AI-enabled tools for teachers is that AI can enable autonomous activity by a computer, and thus when a teacher delegates work to an AI-enabled tool, it may carry on with that work somewhat independently. Professor Inge Molenaar⁴⁰ has wondered about the challenges of control in a hybrid teaching scenario: When should a teacher be in control? What can be delegated to a computational system? How can a teacher monitor the AI system and override its decisions or take back control as necessary? Figure 8 expresses the tension around control. To the left, the teacher is fully in control, and there is no use of AI in the classroom. To the right, the technology is fully in control with no teacher involved—a scenario which is rarely desirable. The middle ground is not one dimensional and involves many choices. Molenaar analyzed products

and suggests some possibilities:

- The technology only offers information and recommendations to the teacher.
- The teacher delegates specific types of tasks to the technology, for example, giving feedback on a particular math assignment or sending out reminders to students before an assignment is due.
- The teacher delegates more broadly to the technology, with clear protocols for alerts, for monitoring, and for when the teacher takes back control.

These and other choices need to be debated openly. For example, we may want to define instructional decisions that have different kinds of consequences for a student and be very careful about delegating control over highly consequential decisions (for example, placement in a next course of study or disciplinary referrals). For human in the loop to become more fully realized, AI technologies must allow teacher monitoring, have protocols to signal a teacher when their judgment is needed, and allow for classroom, school, or district overrides when they disagree with an instructional choice for their students. We cannot forget that if a technology allows a teacher choice—which it should—it will take significant time for a teacher to think through and set up all the options, requiring greater time initially.

Challenge: Making Teaching Jobs Easier While Avoiding Surveillance

We also recognize that the very technologies that make jobs easier might also introduce new possibilities for surveillance (Figure 9). In a familiar example, when we enable a voice assistant in the kitchen, it might help us with simple household tasks like setting a cooking timer. And yet the same voice assistant might hear things that we intended to be private. This kind of dilemma will occur in classrooms and for teachers. When they enable an AI-assistant to capture data about what they say, what teaching resources they search for, or other behaviors, the data could be used to personalize resources and recommendations for the teacher. Yet the same data might also be used to monitor the teacher, and that monitoring might have consequences for the teacher. Achieving trustworthy AI that makes teachers' jobs better will be nearly impossible if teachers experience increased surveillance. A related tension is that asking teachers to be "in the loop" could create more work for teachers if not done well, and thus, being in the loop might be in tension with making teaching jobs easier. Also related is the tension between not trusting AI enough (to obtain assistance) or trusting it too much (and incurring surveillance or loss of privacy). For example, researchers have documented that people will follow instructions from a robot during a simulated fire emergency even when (a) they are told the robot is broken and (b) the advice is obviously

wrong.⁴¹ We anticipate teachers will need training and support to understand how and when they will need to exercise human judgment

Challenge: Responding to Students' Strengths While Protecting Their Privacy

Educators seek to tackle inequities in learning, no matter how they manifest locally (e.g. in access to educational opportunities, resources, or supports). In culturally responsive⁴² and culturally sustaining⁴³ approaches, educators design materials to build on the “assets”—individual, community, and cultural strengths that students bring to learning. Along with considering assets, of course, educators must meet students where they are, including both strengths and needs. AI could assist in this process by helping teachers with customizing curricular resources, for example. But to do so, the data inputted in an AI-enabled system would have to provide more information about the students. This information could be, but need not be, demographic details. It could also be information about students’ preferences, outside interests, relationships, or experiences.⁴⁴ What happens to this data, how it is deleted, and who sees it is of huge concern to educators. As educators contemplate using AI-enabled technologies to assist in tackling educational inequities, they must consider whether the information about students shared with or stored in an AI-enabled system is subject to federal or state privacy laws, such as FERPA. Further, educators must consider whether interactions between students and AI systems create records that must be protected by law, such as when a chatbot or automated tutor generates conversational or written guidance to a student. Decisions made by AI technologies, along with explanations of those decisions that are generated by algorithms may also be records that must be protected by law. Therein, a third tension emerges, between more fully representing students and protecting their privacy (Figure 10). Further, representation would be just a start toward a solution. As discussed earlier in this report, AI can introduce algorithmic discrimination through bias in the data, code, or models within AIenhanced edtech. Engineers develop the pattern detection in AI models using existing data, and the data they use may not be representative or may contain associations that run counter to policy goals. Further, engineers shape the automations that AI implements when it recognizes patterns, and the automations may not meet the needs of each student group with a diverse population. The developers of AI are typically less diverse than the populations they serve, and as a consequence, they may not anticipate the ways in which pattern detection and automation may harm a community, group, or individual. AI could help teachers to customize and personalize materials for their students, leveraging the teacher’s understanding of student needs and strengths. It is time

consuming to customize curricular resources, and teachers are already exploring how AI chatbots can help them design additional resources for their students. An elementary school teacher could gain powerful supports for changing the visuals in a storybook to engage their students or for adapting language that poorly fits local manners of speaking or even for modifying plots to incorporate other dimensions of a teacher’s lesson. In the Learning section, we noted that AI could help identify learner strengths. For example, a mathematics teacher may not be aware of ways in which a student is making great sense of graphs and tables about motions when they are in another teacher’s physics classroom and might not realize that using similar graphs about motion could help with their linear function lesson. AI might help teachers when they seek to reflect student strengths by creating or adapting instructional resources. Yet, the broad equity challenges of avoiding algorithmic discrimination while increasing community and cultural responsiveness must be approached within the four foundations we earlier outlined: human in the loop, equity, safety and effectiveness, and evaluation of AI models. We cannot expect AI models to respect cultural responsiveness. The Department is particularly concerned that equity is something that engaged educators and other responsive adults are in the best position to address and something that is never solely addressable as a computational problem.

Questions Worth Asking About AI for Teaching

As leaders in both pre-service and post-service teacher education contemplate how AI can improve teaching (along with policymakers, developers, and researchers), we urge all in the ecosystem to spend more time asking these questions:

- Is AI improving the quality of an educator’s day-to-day work? Are teachers experiencing less burden and more ability to focus and effectively teach their students?
- As AI reduces one type of teaching burden, are we preventing new responsibilities or additional workloads being shifted and assigned to teachers in a manner that negates the potential benefits of AI?
- Is classroom AI use providing teachers with more detailed insights into their students and their strengths while protecting their privacy?
- Do teachers have oversight of AI systems used with their learners? Are they exercising control in the use of AI-enabled tools and systems appropriately or inappropriately yielding decision-making to these systems and tools?
- When AI systems are being used to support teachers or to enhance instruction, are the protections against surveillance adequate?
- To what extent are teachers able to exercise voice and decision-

making to improve equity, reduce bias, and increase cultural responsiveness in the use of AI-enabled tools and systems?

Key Recommendation: Inspectable, Explainable, Overridable AI

In the Introduction, we discuss the notion that when AI is incorporated into a system, the core of the AI is a model. In the Learning section, we discuss that we need to be careful that models align to the learning we envision (e.g., that they aren't too narrow). Now, based on the needs of teachers (as well as students and their families/caregivers), we add another layer to our criteria for good AI models: the need for explainability.⁴⁵ Some AI models can recognize patterns in the world and do the right action, but they cannot explain why (e.g., how they arrived at the connection between the pattern and the action). This lack of explainability will not suffice for teaching; teachers will need to know how an AI model analyzed the work of one of their students and why the AI model recommended a particular tutorial, resource, or next step to the student. Thus, explainability of an AI system's decision is key to a teacher's ability to judge that automated decision. Such explainability helps teachers to develop appropriate levels of trust and distrust in AI, particularly to know where the AI model tends to make poor decisions. Explainability is also key to a teacher's ability to monitor when an AI system may be unfairly acting on the wrong information (and thus may be biased. We discuss bias and fairness more in the Assessment section next). Surrounding the idea of explainability is the need for teachers to be able to inspect what an AI model is doing. For example, what kinds of instructional recommendations are being made and to which students? Which students are being assigned remedial work in a never ended loop? Which are making progress? Dashboards in current products present some of this information, but with AI, teachers may want to further explore which decisions are being made and for whom and know of the student-specific factors that an AI model had available (and possibly which factors were influential) when reaching a particular decision. For example, some of today's adaptive classroom products use limited recommendation models that only consider student success on the last three mathematics problems and do not consider other variables that a teacher would know to consider, such as whether a student has an IEP Plan or other needs. Our call for attending to equity considerations as we evaluate AI models requires information about how discriminatory bias may arise in particular AI systems and what developers have done to address it. This can only be achieved with transparency for how the tools use datasets to achieve outcomes and what data they have available or that a teacher could include in her judgement but are not available to the system

(IEP status is offered as an example above). Teachers will also need the ability to view and make their own judgement about automated decisions, such as decisions about which set of mathematics problems a student should work on next. They need to be able to intervene and override decisions when they disagree with the logic behind an instructional recommendation.⁴⁶ Teachers need protection against adverse ramifications when they assert human judgement over an AI system's decision.

Formative Assessment

Formative assessment is traditionally a key use of edtech because feedback loops are vital to improving teaching and learning.⁴⁷ As we have emphasized throughout this report, a top priority with AI is to keep humans in the loop and in control, which includes focusing on the people engaged with formative assessments: students, teachers, school leaders, families/caregivers, and others who support learners. In the definition below, please note the overlap between definitions of AI and formative assessment; both have to do with detecting patterns and choosing a future course of action (that adapts to learner strengths and needs).

Building on Best Practices

A number of dimensions hold potential for shaping the future of formative assessments, and many have ready extensions to the field of AI-enabled systems and tools. For example, the 2017 NETP discussed how technology can lead to improved formative assessments along seven dimensions, listed below:

1. Enabling Enhanced Question Types: to give students more ways to show what they know and can do.
2. Measurement of Complex Competencies: to better elicit growth in important skills that go beyond typical subject matter standards, for example, in measuring practices, social skills like teamwork, self-regulation, and work-relevant skills (e.g., making presentations or leading teams).
3. Providing Real-Time Feedback: to maintain and increase student engagement and to support effective learning, providing timely and helpful responses and suggestions to each learner.
4. Increasing Accessibility: to include neurodiverse learners and to engage learners' best communication capabilities as they share what they know and can do.
5. Adapting to Learner Ability and Knowledge: to make assessments more precise and efficient.
6. Embedded Assessment in the Learning Process: to emphasize an assessment's role in improving teaching and learning (this report does not focus on assessment for accountability purposes).
- 7.

Assess for Ongoing Learning: to reveal progress over time and not just predetermined milestones. AI models and AI-enabled systems may have potential to strengthen formative assessments. In one example, a question type that invites students to draw a graph or create a model can be analyzed with AI algorithms,⁴⁹ and similar student models might be grouped for the teacher to interpret. Enhanced formative assessment may enable teachers to better respond to students' understanding of a concept like "rate of change" in a complex, real-world situation. AI can also give learners feedback on complex skills, such as learning American Sign Language⁵⁰ or speaking a foreign language⁵¹ and in other practice situations where no person is available to provide immediate feedback. Generally, an AI assistant may be able to reduce the load for teachers related to grading simpler aspects of student responses, allowing the teacher to focus their specialized judgment on important qualities of a whole essay or a complex project. We also may be able to better provide feedback with accessibility. For example, an AI-enabled learning technology may be able to interact verbally with a student about their response to an essay prompt, asking questions that guide the student to clarify their argument without requiring the student to read a screen or type at a keyboard. In the examples shared earlier in the Learning section, we also see that AI can be embedded in the learning process, providing feedback to students as they work to solve a problem, rather than only later after the student has reached a wrong answer. When formative assessment is more embedded, it can better support learning, and timely feedback is critical.⁵² Although there are many points of connection like these between AI and formative assessments, our listening sessions also revealed attendees' desire to tackle some existing shortcomings in the field of formative assessment; namely, the time-consuming and sometime onerous nature of taking tests, quizzes, or other assessments and the lack of perceived value in the feedback loop by teachers and students.

Implications for Teaching and Learning

Real-time instructional feedback can be beneficial when it helps learners and teachers to improve. But common experience too often leaves students and teachers with unpleasant feelings toward assessment and thus poses a provocative conflict between the potential benefits of data collected through formative assessments and the practical implications of administering additional assessments in classrooms and schools. Some AI-enabled systems and tools seek to address this potential conflict. For example, one AI-enabled reading tutor listens to students as they read aloud and provides on-the-spot feedback to improve their reading.⁵³ Students reportedly enjoyed

reading aloud, and the approach was effective. Researchers have also embedded formative assessments in games so that students can show how well they understand Newtonian physics as they play increasingly difficult levels of a game.⁵⁴ If a student can more easily ask for and receive help when they feel frustrated or confused, reducing those feelings can feel encouraging. Student feelings of safety, confidence, and trust in the feedback generated by these AI-enabled systems and tools are essential to showcase their learning. That focus on learning growth and gains is optimal (absent negative consequences or a high-stakes environment).⁵⁵ AI-enhanced formative assessments may have the potential to save teachers' time (e.g., time spent on grading), allowing the instructor to spend more time engaged in helping students. AIenhanced assessments may also benefit teachers if they provide detailed insights about student strengths or needs that may not be visible and if they support instructional adaptation or improvement by suggesting a small set of evidence-based recommendations for helping students master content. Such assessments may also be helpful outside of the classroom if it can provide feedback when the teacher is not available, for example, in completing homework or practicing a concept during study hall. As we discussed in the Teaching section, an essential aspect of deploying AI-based formative assessment must be centering teachers in system design.

Insight: AI Can Enhance Feedback Loops

The term “formative assessment” does not singularly mean a test or a measurement. Assessment becomes formative when it results in useful reflections and changes to the course of teaching, learning, or both.⁵⁶ The term “feedback loops” emphasizes that measurement is only part of the process. Feedback loops that lead to instructional improvement—including adaptations in teaching and learning—yield the strongest outcomes for students. We also use “feedback loops” as a plural term because there are many types and levels of loops that are important. Students can benefit from feedback when they work individually, as a member of a small group, or in a classroom discussion. Feedback loops are valuable “in the moment”—for example, as a student practices a skill. Further, feedback loops are valuable when they cover larger spans of effort and reflections, such as at the end of presenting a project or term paper. In addition, feedback loops can assist teachers, for example, helping them notice continuous improvement of products and the implementation of programs. Due to the importance of feedback loops, formative assessment could be a leading area for schools’ explorations of powerful uses of AI in teaching and learning. Educators can build upon alignments between their long-standing visions for formative assessment and the emerging

capabilities that AI holds. Further, the professional assessment community brings a toolkit for asking and answering questions about topics like bias and fairness. The psychometric toolkit of methods is a strong start toward the questions that must be asked and answered because it already contains ways to measure bias and fairness and, more generally, to benchmark the quality of formative assessments. But as our discussion reveals, AI can only make feedback loops better if we keep a firm eye on the weaknesses of AI and how AI introduces new concerns.

An Example: Automated Essay Scoring

One instructive example is Automated Essay Scoring (AES). To become strong writers, which is a valuable life skill, students need regular and specific feedback. However, reviewing and providing feedback on essays is very time consuming for humans. Hence, Ellis Page provided a first vision for computer programs that could review and provide feedback on student essays in 1966⁵⁷, and much effort has gone into AES technologies in the intervening 56 years. Many research review articles are available to summarize the progress, which has been impressive.⁵⁸ Further, some of today's applications of AES technologies will be familiar to readers, such as Grammarly, Turnitin, and the various essay analysis engines used by publishers and assessment companies. Also note that while the traditional AES functionality emphasizes scoring or rating essays, newer AI-enabled products focus more on providing students with constructive criticism and developing their skills as writers. Writing is a life skill that is important to the pursuit of college and career ambitions, and developing writers require comprehensive feedback. If developers could inexpensively augment human feedback to developing writers with AI feedback, it's possible that support for learning to write could become more equitable. And yet, AES is an instructive example because researchers have analyzed limitations, too.⁵⁹ AES technologies in AI can analyze some features of student essays but can also be misled by the length of an essay, by a student who places appropriate keywords in sentences that don't make sense, and other flaws that a human reader would easily notice. In a telling quote, one team that reviewed the state of the art wrote this: The authors further note that while human and AI judgements of essays may correlate, people and computers are not noticing the same things in student writing. Due to these limitations, we must continue to emphasize a human in the loop foundation for AI-enhanced formative assessment. AI may support but not replace high-quality, human-led processes and practices of formative assessment in schools.

Key Opportunities for AI in Formative Assessment

Based on the listening sessions we held, we see three key areas of opportunity in supporting formative assessment using AI systems and models. First, we recommend a strong focus on measuring what matters⁶¹ and particularly those things that have not been easily measured before and that many constituents would like to include in feedback loops. The example above, AES, was chosen because writing remains a valuable academic, workplace, and life skill. Looking at community goals through the lens of their visions for their high school graduates, we see that families/caregivers, students, and community leaders want to nurture graduates who solve problems adaptively, who communicate and collaborate well, who persevere and self-regulate when they experience challenges. “What matters” today reaches beyond a sole focus on the core academic content measured by large-scale summative assessments, to support students and teachers with actionable feedback that nurtures the broader skills students need to succeed and thrive. Further, within core academic content, AI may help us to provide feedback on the more realistic and complex aspects of doing math, for example, or investigating scientific phenomena, understanding history, or discussing literature. Second, we’d like to see a strong focus on improving help-seeking and help-giving.⁶² Asking for and giving help is crucial to learning⁶³ and practicing a growth-mindset and central to the notion of human feedback loops. Students may not always know when they need help. In one example, computer algorithms can detect a student who is “wheel spinning” (working hard on mastering content but not making progress).⁶⁴ A student who is working hard may not feel like they need help, and the teacher may not be aware that the student is struggling if he or she appears to be “on task.” AI may also be helpful by highlighting for students and teachers what forms of assistance have been most useful to the student in the recent past so that an educator can expand access to specific assistance that works for that individual student. Finally, educators may learn things from AI-enabled systems and tools that give feedback and hints during the completion of homework, utilizing that feedback to later reinforce concepts in direct instruction and strengthen the one-on-one support provided to students.⁶⁵ AI-enabled systems and tools can provide teachers with additional information about the students’ recent work, so their instructor has a greater contextual sense as they begin to provide help. Third, we advocate for teachers and students to be strongly involved in designing feedback loops as developers produce AI-enhanced formative assessments so they can directly voice what would make assessments less onerous and more convenient and valuable to them.⁶⁶ Earlier in the

Teaching section, we emphasized how important it is to involve teachers in designing, selecting, and evaluating AI-enhanced technologies. Students need to be centered, too. They are experiencing AI in their everyday lives, and they have strong opinions on what is valuable and safe. There are local and cultural variations in how people provide and receive feedback, so adjusting feedback to align with community norms is important.

Key Recommendation: Harness Assessment Expertise to Reduce Bias

Bias and fairness are important issues in assessment design and administration,⁶⁷ and they hold relevance for the area of AI-enabled assessment. In traditional assessment, a test item might be biased if unnecessary details are included that differentially advantage some students (e.g., a story-based item that references a sport that only boys play regularly may be less helpful to girls). As discussed earlier, with AI, we now must worry about algorithmic discrimination which can arise due to the manner in which AI algorithms are developed and improved from large datasets of parameters and values that may not represent all cohorts of learners. Algorithmic discrimination is not just about the measurement side of formative assessment; it is also about the feedback loop and the instructional interventions and supports that may be undertaken in response to data collected by formative assessments. There is a question both about access to such interventions and the quality or appropriateness of such interventions or supports. When an algorithm suggests hints, next steps, or resources to a student, we have to check whether the help-giving is unfair because one group systematically does not get useful help which is discriminatory. Fairness goes beyond bias as well. In AI-enabled formative assessment, both the opportunity to learn through feedback loops, as well as the quality of learning in and outside of such loops, should be addressed. Issues of bias and fairness have arisen in traditional assessments, and the field of psychometrics has already developed valuable tools to challenge and address these issues.⁶⁸ Assessment as a field may have a head start on tackling bias and fairness for AI in education. And yet the issues expand with AI, so the work is not done. Strong and deliberate attention to bias and fairness is needed as future formative assessments are developed.

Related Questions

As indicated, formative assessment is an area in which AI is expanding along a continuum that can be guided by visions already in place, such as the 2017 NETP. It is an area in which AI is poised to grow, especially with capabilities that power more feedback loops in student learning. As this growth takes place, we suggest ongoing attention to the following questions:

- Is formative

assessment bringing benefits to the student learning experience and to the efficacy of classroom instruction? • Are humans being centered in AI-enabled formative assessment and feedback loops? • Are we providing empowering professional development to teachers so they can leverage feedback loops and safeguard against concerns? • To what extent are the developers and implementers of AI-enabled systems and tools tackling new sources of algorithmic bias and continuing to make assessment fairer? • Are governance policies regarding who owns, controls, and can view or use AI-enabled formative assessment data appropriate and adequate? • Do we have sufficient guardrails against misuse of formative assessment data or automatically generated interpretations of student achievement and learning, such as on dashboards? • Is trust in an AI-enabled assessment system, feedback loops, and data generated by such assessments growing or diminishing?

Research and Development

Policy relies upon research-based knowledge; likewise, improving practice depends on feedback loops that analyze empirical evidence. Consequently, the 2010 NETP specified a series of “grand challenges” which were “R&D problems that might be funded and coordinated at a national level.” One 2010 NETP grand challenge was to create personalized learning systems that continuously improve as they are used: Since 2010, much R&D has addressed this challenge. Conferences about learning analytics, educational data mining, and learning at scale have blossomed. Developers have created platforms that use algorithms and the analysis of big data to tune learning experiences. The challenge has not been fully achieved, and further work on this challenge is still relevant today.

Insight: Research Can Strengthen the Role of Context in AI

Despite the relevance of 2010’s grand challenges, it has become apparent that the R&D community is now looking to expand their attention. The 2010 challenges were stated as technical problems. Today’s researchers want to more deeply investigate context, and today’s tech companies want to develop platforms that are responsive to the learners’ characteristics and situations more broadly—not just in terms of narrow cognitive attributes. We see a push to transform R&D to address context sensitivity. We look forward to new meanings of “adaptive” that broaden outward from what the term has meant in the past decade. For example, “adaptive” should not always be a synonym of “individualized” because people are social learners. Researchers therefore are broadening

“adaptivity” to include support for what students do as they learn in groups, a form of learning that is prevalent in schools across the U.S. The focus on context is not an accident. Context is a traditional challenge in AI.⁷⁰ Thus, researchers and developers are wise to prioritizing context. Unless we invest more in AI that is context-sensitive, it is quite likely that AI will break and fail to achieve educational goals. Agreeing to prioritize context won’t be easy. As illustrated above in Figure 12, there will be a tension between depth of context and pace of technological advances in AI R&D. On the one hand, AI is sometimes presented as a race to be the first to advance new techniques or scale new applications—innovation is sometimes portrayed as rapidly going to scale with a minimally viable product, failing fast, and only after failure, dealing with context. On the other hand, researchers and developers see that achieving good innovations with AI in education will clearly require bringing more context into the process early and often. For example, researchers highlight that humans must be continually adjusting the goals for technology and have noted that when we set forth goals, we often don’t yet fully understand context; and as we learn about context, the goals must change.⁷¹ This suggests that context must be prioritized early and habitually in R&D; we don’t want to win a race to the wrong finish line. Further, intensifying focus on context in this work will change the nature of the R&D. There won’t be just one type of change in R&D because context has multiple meanings. Attendees in our listening sessions described four types of context necessary for the future. We list these four types of context below and then expand on each one in its own section. These four types emerged as topics of provocations to think differently about R&D but certainly do not exhaust the important ways of investigating context.

1. Focus on the Long Tail: How could we use big data and AI to pay more attention to the “long tail” of edtech use—going beyond a few “most typical” ways of using emerging technology and instead solving for digital equity and inclusion?
2. Partnership in Design-Based Research: How can we change who is involved and influential in designing the future of AI in education to more centrally include students, teachers, and other educational constituents?
3. Connect with Public Policy: How can work on AI in education build on general advances in AI ethics, safety, and regulation and contribute additional advances specific to educational policy?
4. Rethink Teacher Professional Development: How can we solve for new systems of teacher professional development (both pre-service and in-service) that align to the increasingly core role of technology in the teaching profession?

Attention to the Long Tail of Learner Variability

At the core of R&D of AI in education, innovators will be building models that fit available data. The increasing scale and prevalence of technologies means that the data is coming from and including a wide range of different contexts and varied ways that people in those contexts engage in teaching and learning. Researchers in our listening sessions drew attention to the promise of AI for addressing “context” by reference to the long tail of learner variability As depicted in Figure 13, learners vary in their strengths and needs. The most frequently occurring mix of strength and needs (also known as “teaching to the middle”) is depicted leftmost, with less frequently occurring mixes spreading to the right. Rising upward, the figure depicts the number of learners who benefit from a particular learning design, pathway, or approach. We argue that AI can bring opportunities to address a wider spectrum of strengths and needs but only if developers and innovators focus on the long tail and not only “teaching to the middle.” For the sake of argument, the figure indicates three zones. In a first zone, curricular resources are mostly standardized, with perhaps a dimension or two of adaptivity. For example, many existing products adapt based on the correctness of student answers and may also provide options to read or hear text in a second language. However, the core of the instructional approach is highly standardized. In a second zone, there is greater balance between how much standardization and how much adaptivity students can access. Universal Design for Learning (UDL) is one set of recommendations for providing learning opportunities in multiple formats and for accommodating different learning progressions.⁷² UDL can enable accommodating more ways in which learners vary, and as teachers know, there are many more important ways to adapt to students than found in today’s edtech products. Students are neurodiverse. They bring different assets from their experiences at home, in their communities, and in their cultures. They have different interests and motivations. And they learn in varied settings—classrooms and schools differ, and at-home students learn in informal settings in ways that could complement school learning. These are all important dimensions of “context.” Zone 3 indicates highly adaptive learning, where standardization is less successful and where we need to discover a wider variety of approaches to engage learners and sustain powerful learning. Researchers in our listening sessions noted the promise of Zone 3 because AI’s ability to recognize patterns in data can extend beyond the most common patterns and because AI’s ability to generate customized content can extend beyond what people can reasonably generate on their own. Notice that although the Zone 1 bar appears to be the tallest, and thus tends to attract initial attention, there

are more students in Zones 2 and 3, the regions where AI can provide more help. Thus, it's important to ask where AI researchers and developers are directing their attention. When we say a model "fits," are we saying it fits the most common and typical uses by teachers and learners? This sort of R&D is easier to do. However, machine learning and AI also can tailor a model to the less common and more culturally specific contexts, too. Therefore, how can constituents cultivate interdisciplinary expertise to direct attention among researchers and developers to focus on the long tail? If we do, the quality of what we do for those represented in that tail can be more adaptive and more context-sensitive. And to be most effective, it will require the integration of contextual, content, and technical expertise. Within the long-tail challenge, the community is wondering how we can get to research insights that are both general and specific enough. When research produces very general abstractions about learning, it often doesn't give developers enough guidance on exactly how to adjust their learning environments. Conversely, when research produces a specific adaptive algorithm that works on one educational platform, it often remains hard to apply to additional platforms; research can be too detailed as well. The research community is also thinking about new partnerships that could bring more data and more diverse perspectives to the table, the topic of the next section. Focusing on the long tail of learner variability is particularly important to addressing a longstanding key research question: "Do new AI-enhanced approaches work to improve learning, and for whom and under what conditions?"

Partnership in Design-Based Research

Of course, teachers must be included in rethinking their own professional development. This thought leads to another priority aspect of context: partnership in design-based research. With regard to inclusive design, attendees in our listening sessions brought up a variety of co-design⁷³ and other participatory processes and goals that can be used in R&D.⁷⁴ By co-design, they mean sharing power with non-researchers and non-developers through all the phases of design and development, which would result in more influence by teachers, students, and other constituents in the shape of AI-enabled edtech. The shift toward co-design was palpable throughout our listening sessions, but as researchers and developers have not standardized on one particular codesign method, we share some representative examples. • Youth can powerfully participate in design when researcher methods include participant co-design. Such research can investigate how to improve edtech while educating students. A listening session attendee asked about developing students' awareness of what data are being collected and how data are being used by developers.

- There is a near future need to go beyond representation so that co-designed solutions consider more generous contexts for broader possibilities, according to attendees.
- The shift of power dynamics is another research-worthy interest of the panel and attendees to understand the balance between a teacher's agency and a machine's suggestions.
- Likewise, such longitudinal research will require both the infrastructure and institutional support to fund necessary experimentation and requisite failures to elicit positive results and safe innovation.
- There is a desire for rapid cycle evaluations with inclusive feedback loops that return to the educators themselves as essential relative to traditional research approaches.
- Many researchers also mentioned a focus on explainable AI as essential to enable participation in the design and evaluation of emerging AI approaches in education. The conversations raised this question: how can co-design provide an empowering form of participation in design and thus achieve digital inclusion goals? Such digital inclusion can span many layers of design, including diverse representation in design of policies around data, design of adaptivity, and other user experiences in AI systems, design of plans for cultivating AI literacy for users of new platforms, and lastly, the design of plans to evaluate systems.

Re-thinking Teacher Professional Development

With regard to teachers as professionals, both researchers and other educators attending our listening sessions were highly concerned about the disconnect between how teachers are prepared versus how they are expected to work with emerging technology. When we discuss learning, teachers are central actors, and thus the contexts in which they are prepared is centrally important to their ability to do great work in current and emerging technological environments. Teacher professional development, professional learning, and leadership (PD or PL) for emerging technologies was seen as an area needing intense re-thinking, and research could lead the way. Today, few who prepare to become a teacher in an established pre-service program learn about the effective use of educational technology in schools and classrooms; those who do have the opportunity to investigate technology rarely think about the structures that shape its use in the classroom and in educational leadership. Consequently, a troubling dichotomy arises between a small set of investigators who specifically consider educational technology in their research on teaching and a broader group of educators who see educational technology as a generic instructional resource. The challenge is high because teacher professional development will remain highly varied by local contexts. Yet insufficient attention to teachers as leaders in the use and

further development of effective educational technology is widespread in teacher professional development research. One response can be in terms of investigating how to nurture greater AI literacy for all teachers. AI literacy is not only important to protect educators and students from possible dangers but also valuable to support teachers to harness the good and do so in innovative ways. A panelist reminded the group that this work implies how we prepare educators with a baseline AI literacy and understanding. More transparency and authentic dialogue can foster trust, which was mentioned by a researcher as a chief concern for all teachers and students. This is not to suggest that AI literacy is a complete or even a simple fix. Researchers want to ask fundamental questions about what it means for teachers to be professionals, especially as emerging technologies gain ground in schools and classrooms—our teachers' professional workplaces. Researchers want to broadly reconceptualize teacher professionalism and to stop treating technology as an add-on element of professional development.

Connecting with Public Policy

Defining human-centered AI for education requires the embrace of a human-centered principle and foundation for developing and formulating policies that govern the application and use of AI more generally throughout society. For example, power dynamics that arise between companies and consumers in society around issues like data ownership will also arise in the education-specific ecosystem. Further, the public discourse in which people are discussing ethics, bias, responsibility, and many other necessary concepts will be happening simultaneously in public policy and in educational ecosystems. One clear implication in our listening sessions was that efforts to improve AI literacy in education could be important and helpful to society more generally. For example, one panelist said that an overarching goal of improving AI literacy is necessary if they are to contribute to how those technologies are designed. Another researcher was interested in how edtech can provide environments where students can experience having difficult discussions across perspectives, an issue which is endemic to present society. A third researcher noted the insufficiencies of prior efforts to contend with algorithmic bias, ethics, and inclusion due to a classroom's complex social dynamics. Researchers want to take a lead in going beyond checkbox approaches to take these issues seriously. And they also acknowledge that engaging with policy is often a new form of context for edtech and AI researchers, many of whom don't have long experiences in policy arenas. Likewise, developers often do have experience with some policy issues, such as data privacy and security, but are now needing to become part of new conversations

about ethics, bias, transparency, and more, a problem that the EdSAFE AI Alliance is addressing through multisector working groups and policy advocacy.⁷⁵

Key Recommendation: Focus R&D on Addressing Context

Attendees who have participated in listening sessions leading up to this report were exceptionally clear that their view of future R&D involved a shift from narrow technical questions to richer contextual questions. This expansive shift toward context, as detailed below, is the foundational orientation that the listening session attendees saw as being necessary to advancing R&D. Attendees included these as dimensions of context:

- learner variability, e.g., in disabilities, languages spoken, and other relevant characteristics;
- interactions with peers, teachers, and others in the learning settings;
- relationships across home, school, and community settings, including cultural assets;
- instructional resources available while learning;
- teacher preparation;
- and
- policies and systems that structure teaching and learning.

To more fully represent the context of teaching and learning, including these and other dimensions of context, researchers will have to work in partnership with others to understand which aspects of context are most relevant to teaching and learning and how they can be usefully incorporated into AI models.

Ongoing Questions for Researchers

As mentioned earlier, people are good at context; AI—not so much. R&D investment in contextrich edtech thus could serve multiple national interests because finding ways to do a better job with context would be a fundamental advancement in AI. Indeed, questions like these reverberate across all applications of AI in society, and education is a centrally good context for investigating them:

- Are AI systems moving beyond the tall portions of the “long tail” to adapt to a greater range of conditions, factors, and variations in how people learn?
- To what extent are AI technologies enhancing rather than replacing human control and judgment of student learning?
- How will users understand the legal and ethical implications of sharing data with AI enabled technologies and how to mitigate privacy risks?
- To what extent does technology account for the complex social dynamics of how people work and learn together, or is technology leading humans to narrow or oversimplify?
- How can we more clearly define what we mean by a context-sensitive technology in terms that are both concrete and broad enough? How can we measure it?
- To what extent are technical indicators and human observations of bias or unfairness working together with human observations?
- How can concerns about ethics and equity in AI technologies become actionable both in R&D, and later, when AI is widely used?
- Are we learning for whom and under

what conditions AI systems produce desired benefits and impacts and avoid undesirable discrimination, bias, or negative outcomes?

Desired National R&D Objectives

Attendees sought immediate progress on some key R&D issues, such as these:

- Clarifying and achieving a consensus on the terms that go beyond data privacy and data security, including ideas like human-centered, value-sensitive, responsible, ethical, and safe so constituents can advocate for their needs meaningfully and consistently
- Creating and studying effective programs for AI literacy for students, teachers, and educational constituents in general, including literacy with regard to the ethics and equity issues specific to AI in educational settings
- Advancing research and development to increase fairness, accountability, transparency, and safety in AI systems used in educational settings
- Defining participatory or co-designed research processes that include educators in the development and conduct of research related to the development, use, and efficacy of AI-enabled systems and tools
- Highlighting and advancing R&D efforts that empower the participation and voices of youth regarding research, data, and design of AI applications for teaching and learning

Longer term desires for a national R&D program include some of the following objectives:

- Funding sustainable partnerships that uncover what context means and how it can be addressed over longer periods of time
- Better connecting goals for “broadening participation” (for example, in STEM learning pathways) to strategies for addressing learner variability and diversity
- Prioritizing research to revitalize support for instructors in light of the increasingly technological nature of K-12, higher education, and workplace learning settings
- Creating infrastructure and new ways of working together beyond individual field-initiated grants so that R&D with big data and leveraging emerging AI capabilities becomes safer and more productive

Recommendations

Earlier, we asked two guiding questions: 1. What is our collective vision of a desirable and achievable educational system that leverages automation while protecting and centering human agency? 2. On what timeline will we be ready with necessary guidelines and guardrails along with convincing evidence of positive impacts, so that we can ethically and equitably implement this vision widely? Answers to the first question are provided throughout the Learning, Teaching,

Assessment, and Research sections. This section turns to a call to action to education leaders and to recommendations. Core to the Department’s perspective is that education will need leadership specific to our sector. Leadership should recognize and build on prior accomplishments in edtech (such as strong prior work on student privacy and school data security) as well as broad frameworks for safe AI (such as the Blueprint for an AI Bill of Rights). Leadership must also reach beyond these accomplishments and frameworks to address emerging opportunities and risks that are specific to novel capabilities and uses of AI in education.

Insight: Aligning AI to Policy Objectives

Individual sections of this policy report provided insights in each of four areas—learning, teaching, assessment, and research. These insights, synthesized from extensive stakeholder consultation and listening sessions, show that the advances in AI can bring opportunities to advance the Department’s policy objectives:

- In support of our objective of attracting and retaining teachers, our nation could focus on AI assistants that make teaching jobs better and provide teachers with the information they need to work closely and empathically with students. An emphasis on teachers in the loop could ensure that AI-enabled classroom technologies keep teachers in the know, in touch with their students, and in control of important instructional decisions. Keeping the teacher in the loop is important to managing risks, as well.
- In support of equitable learning, especially for those most affected by the pandemic, AI could shift edtech from a current deficit-based model to a strengths-based alternative. In addition to finding student weaknesses and assigning fixes, edtech could make recommendations based on strengths that students bring to learning and how adapting to the whole student—a cognitive, social, and self-regulating person—could enable more powerful learning. Adapting to the whole student should include supporting students with disabilities as well as English learners. With regard to equity, we must remain highly attuned to the challenges of bias (which are inherent to how AI systems are developed) and take firm action to ensure fairness.
- With regard to growth trajectories to successful careers, AI-enabled assessments could provide students and teachers with formative guidance on a wider range of valuable skills, focusing on providing information that enhances learning. Aligned with the human-centric view, we should take a systems view of assessments where students, teachers, and others remain at the center of instructional decision making
- With regard to equity, as research advances and brings more context into AI, we will be better able to use AI to support goals that require customization of learning resources, such as enabling teachers to more easily transform

materials to support neurodiverse learners and increase responsiveness to local communities and cultures. Going forward, educational leaders need to bring these and their own policy priorities to the table at every discussion about AI, driving the conversation around human priorities and not only their excitement about what new technology might do. Fundamentally, AI seeks to automate processes that achieve goals, and yet, AI should never set goals. The goals must come from educators' vision of teaching and learning and educators' understanding of students' strengths and needs.

Calling Education Leaders to Action

We summarize seven recommendations for policy action. These recommendations are for education leaders. In the introduction, we note the necessity of involving education constituents in determining policies for AI. We also observed throughout our listening sessions that people coming from many different roles in education all have passion, knowledge, and insights to contribute. In our view, all types of constituents can be education leaders. We are reluctant to suggest any constituent role is more important to advance any of the recommendations, but we call out specific needs for action within some of the recommendations where it is warranted.

Recommendation #1: Emphasize Humans in the Loop

We start with a central recommendation throughout this report. This recommendation was a clear constituent favorite. Indeed, across more than 700 attendees in our listening sessions, the predominant discussion tackled how constituents can achieve a consensus vision for AI-enabled edtech where humans are firmly at the center. The Blueprint for an AI Bill of Rights similarly calls for “access to timely human consideration and remedy by a fallback and escalation process if an automated system fails, it produces an error, or you would like to appeal or contest its impacts...” Building on this consensus, we call upon all constituents to adopt “humans in the loop” as a key criterion for educational use of AI. We envision a technology-enhanced future more like an electric bike and less like robot vacuums. On an electric bike, the human is fully aware and fully in control, but their burden is less, and their effort is multiplied by a complementary technological enhancement. Robot vacuums do their job, freeing the human from involvement or oversight. Although teachers should not be the only humans involved in loops, Figure 5 provided examples of three types of teacher loops that are central to education and can be used to illustrate what “human in the loop” means. Here, we use the example of an AI chatbot to elaborate on the meaning of the loops. First, as students become involved in extended interactions with AI chatbots, teachers

will need to educate students about safe AI use, monitor their use, and provide human recourse when things go astray. Second, teachers are beginning to use chatbots to plan personalized instruction for their students; they will need to be involved in loops with other teachers to understand effective prompts, to know how to analyze AI-generated lesson plans for flaws, and to avoid the human tendency to overly trust AI systems and underapply human judgement. Third, teachers need to be involved in the design and evaluation of AI systems before they are used in classrooms and when needs for improvement are observed. In one example, to design AI-generated homework support for students, teachers' in-depth understanding of the cognitive, motivational, and social supports their students need will provide much-needed guidance as a homework-support chatbot is designed. In framing AI in education, this report advances a key recommendation of "human in the loop" AI because the phrase readily communicates a criterion that everyone can use as they determine which AI-enabled systems and tools are appropriate for use in teaching and learning. In a rather technical field, human in the loop is an approachable and humanistic criterion. Rather than suggesting that AI-enabled systems and tools should replace teachers, this term instead solidifies the central role of educators as instructors and instructional decision makers, while reinforcing the responsibility of teachers to exercise judgement and control over the use of AI in education. It resonates with the important idea of feedback loops, which are highly important to how people teach and learn. It also aligns with the ideas of inspectable, explainable, severable, and overridable AI. The Department agrees with listening session participants who argued that teachers should not be the only humans in the loop and calls upon parents, families, students, policy makers, and system leaders to likewise examine the "loops" for which they are responsible, critically analyze the increasing role of AI in those loops, and determine what they need to do to retain support for the primacy of human judgement in educational systems.

Recommendation #2: Align AI Models to a Shared Vision for Education

As we have discussed across every section of this report, AI technologies are grounded in models, and these models are inevitably incomplete in some way. It is up to humans to name educational goals and measure the degree to which models fit and are useful—or don't fit and might be harmful. Such an assessment of how well certain tools serve educational priorities may seem obvious, but the romance of technology can lead to a "let's see what the tech can do" attitude, which can weaken the focus on goals and cause us to adopt models that fit our priorities poorly. Here we call upon educational policy and decision makers at the local, state, and federal level to

use their power to align priorities, educational strategies, and technology adoption decisions to place the educational needs of students ahead of the excitement about emerging AI capabilities. We want to strengthen their attention to existing state, district, and school-level policies that guide edtech adoption and use, such as the four levels of evidence in ESSA, the privacy requirements of FERPA, and enhanced policies to come. Local education leaders know best what their urgent educational priorities are. Every conversation about AI (or any emerging technology) should start with the educational needs and priorities of students front and center and conclude with a discussion about the evaluation of effectiveness re-centered on those needs and priorities. Equity, of course, is one of those priorities that requires constant attention, especially given the worrisome consequences of potentially biased AI models. We especially call upon leaders to avoid romancing the magic of AI or only focusing on promising applications or outcomes, but instead to interrogate with a critical eye how AI-enabled systems and tools function in the educational environment. We ask leaders to distrust broad claims and ask six types of questions, listed below. Throughout this report, we elaborated on which characteristics of AI model use in education are most important to evaluate for alignment to intended educational goals. To aid leaders, we summarize our insights about AI models and their use in educational tools and systems in Figure 14. In this figure, we center teaching and learning in all considerations about the suitability of an AI model for an educational use. Humans remain in the loop of defining, refining, and using AI models. We highlight the six desirable characteristics of AI models for education (elaborating from principles in the Blueprint for an AI Bill of Rights to fit the specifics of educational systems): 1. Alignment of the AI Model to Educators' Vision for Learning: When choosing to use AI in educational systems, decision makers prioritize educational goals, the fit to all we know about how people learn, and alignment to evidence-based best practices in education. 2. Data Privacy: Ensuring security and privacy of student, teacher, and other human data in AI systems is essential. 3. Notice and Explanation: Educators can inspect edtech to determine whether and how AI is being incorporated within edtech systems. Educators' push for AI models can explain the basis for detecting patterns and/or for making recommendations, and people retain control over these suggestions. 4. Algorithmic Discrimination Protections: Developers and implementers of AI in education take strong steps to minimizing bias and promoting fairness in AI models. 5. Safe and Effective Systems: The use of AI models in education is based on evidence of efficacy (using standards already established in education for this purpose) and work for diverse learners and in

varied educational settings.

6. Human Alternatives, Consideration and Feedback: AI models that support transparent, accountable, and responsible use of AI in education by involving humans in the loop to ensure that educational values and principles are prioritized. Although we first address our recommendation to interrogate how educational systems use AI models to educational leaders who adopt technologies, other leaders also have integral roles to play. Teachers and students, as well as their families/caregivers, contribute significantly to adoption decisions also. And leaders and parents must support educators when they question or override an AI model based on their professional wisdom. Developers of technologies need to be forthcoming about the models they use, and we may need policymakers to create requirements for disclosure so that the marketplace can function on the basis of information about AI models and not only by the claims of their benefits. We also emphasize the need for a government role. AI models are made by people and are only an approximation to reality. Thus, we need policies that require transparency about the AI models that are embedded in educational systems, as well as models that are inspectable, explainable, and overridable. Our listening sessions featured constituent calls for government doing more to hold developers accountable for disclosing the types of AI models they employ in large-scale products and the safeguards included in their systems. Government leaders can make a positive contribution to market conditions that enable building trust as AI systems are procured and implemented in education. We discuss these guidelines more in recommendation #4, which is about building trust.

Recommendation #3: Design Using Modern Learning Principles

We call for the R&D sector to ensure that product designs are based on best and most current principles of teaching and learning. The first decade of adaptivity in edtech drew upon many important principles, for example, around how to sequence learning experiences and how to give students feedback. And yet the underlying conception was often deficit-based. The system focused on what was wrong with the student and chose pre-existing learning resources that might fix that weakness. Going forward, we must harness AI's ability to sense and build upon learner strengths. Likewise, the past decade of approaches was individualistic, and yet we know that humans are fundamentally social and that learning is powerfully social. Going forward, we must build on AI capabilities that connect with principles of collaborative and social learning and which respect the student not just for their cognition but also for the whole human skill set. Going forward, we also must seek to create AI systems that are culturally responsive and culturally sustaining, leveraging

the growth of published techniques for doing so. Further, most early AI systems had few specific supports for students with disabilities and English learners. Going forward, we must ensure that AI-enabled learning resources are intentionally inclusive of these students. The field has yet to develop edtech that builds upon each student’s ability to make choices and to self-regulate in increasingly complex environments. We have to develop edtech that expands students’ abilities to learn in creative modes and to expand their ability to discuss, write, present, and lead. We also call upon educators to reject uses of AI that are based solely on machine learning from data—without triangulation based on learning theory and knowledge from practice. Achieving effective and equitable educational systems requires more than processing “big data,” and although we want to harness insights from data, human interpretation of data remains highly important. We reject a technological determinism in which patterns in data, on their own, tell us what to do. Applications of AI in education must be grounded in established, modern learning principles, the wisdom of educational practitioners, and should leverage the expertise in the educational assessment community around detecting bias and improving fairness.

Recommendation #4: Prioritize Strengthening Trust

Technology can only help us to achieve educational objectives when we trust it. Yet, our listening sessions revealed the ways in which distrust of edtech and AI is commonplace. Constituents distrust emerging technologies for multiple reasons. They may have experienced privacy violations. The user experience may be more burdensome than anticipated. Promised increases in student learning may not be backed by efficacy research. They may have experienced unanticipated consequences. Unexpected costs may arise. Constituents may distrust complexity. Trust needs to incorporate safety, usability, and efficacy. The Department firmly takes the stance that constituents want AI that supports teachers and rejects AI visions that replace teachers. And yet, teachers, students, and their families/caregivers need support to build appropriate levels of trust in systems that affect their work. In the broader ecosystem, trustworthy AI is recognized as a multidimensional problem (including the dimensions of Figure 14, above). If every step forward does not include strong elements of trust building, we worry that distrust will distract from innovation serving the public good that AI could help realize. We expect that associations and societies have a key role in strengthening trust. Some important associations like the State Educational Technology Directors Association and the Consortium for School Network work with edtech leaders, and parallel organizations like EDUCAUSE work with postsecondary leaders.

Other associations and societies work with teachers, education leaders, and education staff developers. Industry networks, like the EdSAFE AI Alliance, can bring together industry leaders to work together to foster trust. Additional societies bring researchers together. These societies and associations have the reach necessary to bring all parts of the educational ecosystem into discussions about trust and also the ability to represent the views of their constituents in cross-cutting policy discussions.

Recommendation #5: Inform and Involve Educators

Our listening sessions also asked for more specific direction on the question of what education leaders should do (see Figure 15). The most frequent responses fit three clusters: the need for guidelines and guardrails, strengthening the role of teachers, and re-focusing research and development. These are activities that constituents are asking for and that could expand trust. The recommendations that follow respond to these requests. In particular, one concern that repeatedly arose in our listening sessions was the potential for AI to result in less respect for educators or less value for their skills. Across the nation, we are now responding to decreasing interest in entering or remaining in the teaching profession. Now is the time to show the respect and value we hold for educators by informing and involving them in every step of the process of designing, developing, testing, improving, adopting, and managing AI-enabled edtech. This includes involving educators in reviewing existing AI-enabled systems, tools, and data use in schools, designing new applications of AI based on teacher input, carrying out pilot evaluations of proposed new instructional tools, collaborating with developers to increase the trustworthiness of the deployed system, and raising issues about risks and unexpected consequences as the system is implemented. We have already seen educators rise to the challenge of creating overall guidelines, designing specific uses of available AI-enabled systems and tools, and ferreting out concerns. And yet, the influence of educators in the future of AI-enabled products cannot be assumed; instead, constituents need policies that put muscle behind it. Could we create a national corps of leading educators representing every state and region to provide leadership? Could we commit to developing necessary professional development supports? Can we find ways to compensate educators so they can be at the forefront of designing the future of education? Our policies should enable educators to be closely involved in design of AI-enabled educational systems. Although we know that the responsibility for informing and involving educators must be distributed at all levels of national and school governance, the Office of Educational Technology can play a key role in

informing and involving educators through its reports, events, outreach, and in a future NETP. Although examples above refer to K-12 teachers, higher education instructors must also be included. We also call on the edtech industry to involve educators throughout their design and development processes. For example, AI-enabled teaching assistants are only likely to help teachers do their job if teachers are thoroughly involved as the assistants are designed. We call upon institutions that prepare teachers to integrate technology more systematically into their programs; for example, the use of technology in teaching and learning should be a core theme across teacher preparation programs, not an issue that arises only in one course.

Recommendation #6: Focus R&D on Addressing Context and Enhancing Trust and Safety

Research that focuses on how AI-enabled systems can adapt to context (including variability among learners) in instructional approaches and across educational settings is essential to answering the question of, “Do specific applications of AI work in education, and if so, for whom and under what conditions?” The italicized phrase points to variability among learners and diversity in the settings for learning. We call upon innovators in R&D to focus their efforts to advance AI on the long tail of learning variability, where large populations of students would benefit from customization of learning. We also call on R&D to lead by establishing how trust can be strengthened in AI-enabled systems, building on the Blueprint’s call for safe and effective systems yet also including education-specific requirements, such as how teachers can be meaningfully involved in design phases, not only in implementation and evaluation. Although many products today are adaptive, some adapt on just one or a few dimensions of variability, such as student’s accuracy in problem solving. As teachers know, there are many more important ways to adapt to students’ strengths and needs. Students are neurodiverse and may have specific disabilities. They bring different assets from their experiences at home, in communities, and in their cultures. They have different interests and motivations. They are in different places in their journeys to master the English language. And they learn in varied settings. Classrooms and schools are different, and at home, students learn in informal settings in ways that could complement school learning. We recommend attention to “context” as a means for expressing the multiple dimensions that must be considered when elaborating the phrase “for whom and under what conditions.” We also acknowledge the role of researchers in conducting evaluations, which must now consider not only efficacy but must also explore where harm may arise and the system problems that can occur through weak trust or over-trust in AI systems. R&D must take the lead

in making AI models more context-sensitive and ensuring that they are effective, safe, and trustworthy for use with varied learners in diverse settings. Although AI has capabilities to find patterns beyond the limited number of variables that people normally think about, AI is not particularly good at understanding and working with context in the ways people do. Over time, we've seen learning sciences grow to be less about individualistic cognitive principles and more encompassing first of social learning and then of the many dimensions of context that matter in learning. Our use of AI needs to follow this trajectory toward context to support educational applications. To achieve human-centric vision, listening session attendees argued that teams will need time and freedom to explore how best to manage the tension between the pace of technological advancement and the need for broader contextual insights—for trust and for safety. They will need time and freedom to pioneer new processes that better involve teachers and students as codesigners, with attention to balancing power dynamics. And they will need to shift attention from older ways of framing priorities (such as achievement gaps) to new ways of prioritizing digital equity. We call on R&D funders to focus resources on the long tail of learner variability, the need for AI-enabled systems that better incorporate context, and time required to get contextual considerations right. We call upon researchers and developers to prioritize challenges of context, trust, and safety in their work to advance AI.

Recommendation #7: Develop Education-Specific Guidelines and Guardrails

Our final recommendation is central to policymakers. A feature of the American educational system is the emphasis on local decision making. With technology growing in complexity at such a rapid pace, it is becoming difficult for local leaders to make informed decisions about the deployment of artificial intelligence. As we have discussed, the issues are not only data privacy and security but extend to new topics such as bias, transparency, and accountability. It will be harder to evaluate promising edtech platforms that rely on AI systems against this evolving, complex set of criteria. Regulations related to key student and family data privacy laws like the Family Educational Rights & Privacy Act (FERPA), the Children's Internet Privacy Act (CIPA), and the Children's Online Privacy Protection Act (COPPA) warrant review and further consideration in light of new and emerging technologies in schools. Laws such as the Individuals with Disabilities Education Act (IDEA) may likewise be considered as new situations arise in the use of AI-enabled learning technologies. As discussed throughout this document, the Blueprint for an AI Bill of Rights is an important framework throughout this work. The Department encourages

parallel work by constituents in all levels of the educational system. In addition to the key federal laws cited immediately above, many states have also passed privacy laws that govern the use of educational technology and edtech platforms in classrooms. Further constituents can expect general frameworks for responsible AI in parallel sectors like health, safety, and consumer products to be informative but not sufficient for education's specific needs. Leaders at every level need awareness of how this work reaches beyond implications for privacy and security (e.g., to include awareness of potential bias and unfairness), and they need preparation to effectively confront the next level of issues.

Next Steps

We are heartened to see intensifying discussions throughout the educational ecosystem about the role of AI. We see progress that we can build upon occurring, as constituents discuss these three types of questions: What are the most significant opportunities and risks? How can we achieve trustworthy educational AI? How can we understand the models at the heart of applications of AI and ensure they have the qualities that align to educational aspirations? The Department developed this report with awareness of contributions arising from many types of organizations and collectives. Internationally, we recognize parallel efforts to consider AI in the European Union, at the United Nations, and indeed throughout the world. We are aware of progress being led by organizations such as UNESCO, the EdSAFE AI Alliance, and research organizations in many countries. We plan to continue cross-agency work, for example, by continuing to coordinate with the Office of Science and Technology Policy and other Federal agencies as agencies implement next steps guided by the Blueprint for an AI Bill of Rights. We see a broad and fertile context for necessary next steps:

- Working within this context and with others, the Department will consider specific policies and regulations so that educators can realize the opportunities of AI in edtech while minimizing risks. For example, the Department is developing a set of AI usage scenarios to strengthen the process of evaluating and enhancing policies and regulations. The principles and practices in the Blueprint for an AI Bill of Rights will be used to ensure the scenarios mitigate important risks and harms.
- Working with constituents (including education leaders; teachers, faculty, support staff, and other educators; researchers; policymakers; funders; technology developers; community members and organizations; and above all, learners and their families/caregivers), we will develop additional resources and events to increase understanding of AI and to involve those who will be most affected by these new technologies.
- Working across

sectors, such as education, innovation, research, and policy, we will revise and update the NETP to guide all constituents toward safe, equitable, and effective AI in education in the United States, in alignment with our overall educational priorities.

31. Guidelines for the Use of Artificial Intelligence in University Courses

Guidelines for the Use of Artificial Intelligence in University Courses Prepared by Juan David Gutiérrez, Universidad del Rosario1 Version 4.3 (21/02/2023). License C.C. BY 4.0

1. Justification:

- We use artificial intelligence tools daily. Consciously or unconsciously, practically all of us use some type of artificial intelligence (AI) in our daily lives. We even use systems that help us perform various university activities. For example: when we are writing a text message or an email and the ‘autocomplete’ tool suggests how to finish writing the sentence; when we search for information through Internet search engines that recommend keywords for the search; when we use word processing software that suggests grammatical corrections; and, when we use automated translation tools, among others. In addition, there are web and mobile AI tools that are freely accessible to anyone with an Internet connection and that allow us to generate text and audio-visual content.
- Some AI can support learning processes and professional activities. There are appropriate uses of such technologies which can be useful for learning and teaching at universities. Moreover, there are different uses for these tools at the professional level; for example, AI tools contribute to activities in all the stages of public policy processes.
- But certain uses of AI are risky, so their use must be informed, transparent, ethical, and responsible. AI tools are not suitable for every type of activity and certain types of uses can be counterproductive to the pedagogical process. Some uses of AI may generate risks for users and third parties. Furthermore, as explained in section 4 of the guidelines, it is essential to be aware of the ethical, environmental, and human rights implications associated with the use of these tools. It is precisely because of these risks that some universities have published recommendations and guidelines on the use of AI in academic contexts.

2. Objectives:

1. Prevention. To prevent situations in which students consciously or unconsciously engage in academic dishonesty.
2. Contribute to digital literacy. Contribute to learning basic knowledge about the use of emerging technologies such as AI, their benefits and risks, and their implications for society.
3. Promotion of responsible use. Promote the responsible and ethical use of emerging technologies, such as AI, in learning processes and for future use in professional life.

3. Rules for the use of AI in and out of the classroom:

1. General rule. In this course, the use of AI as a support tool to carry out different learning activities is allowed. The parameters for the use of these tools described below distinguish between ‘low risk’ and ‘high risk’ tools, according to the risks that the respective AI generates for the pedagogical process, for the users, and for third parties.
2. Use of low-risk AI. You may freely use low-risk AI tools, i.e., those that allow you to correct or review student-generated content or those that allow you to collect and process data. For example, grammar correction tools, translation tools, audio-to-text transcription tools, and Internet information search tools, among others. It is recommended that where such tools enable you to do meaningful work (e.g., translation), you indicate their use in the appropriate section (e.g., methodology section).
3. Use of high-risk AI. AI tools that generate content (text and/or audio-visual), i.e., textgenerating AI (generative AI, such as ChatGPT) and stable diffusion AI for image generation (Stable Diffusion, such as DALL-E 2 or Midjourney) are considered high-risk in these guidelines. You may only include AI-generated content in your individual and group work when four requirements are fulfilled:
 - a. Informed use. Prior to using the tool, research who or what company developed the tool, how it was developed, how it works, what functions it can perform, and what limitations and/or risks it presents.
 - b. Transparent use. In your work, indicate in detail and expressly which tool you used and how you used it (a requirement currently requested by scientific journals to those who submit manuscripts for peer review).
 - c. Ethical use. The manuscripts must distinguish what was written or produced directly by you and what was generated by an AI tool. On this point,

general citation rules apply, e.g., use quotation marks if you include textual paragraphs. Violations of this policy, particularly with respect to the second requirement, will be considered an infringement to academic integrity. In addition, ChatGPT is not a person, so it cannot be considered your co-author (just as you should not include Google as your co-author). Some scientific journals have already updated their publication policies to clarify that language models such as ChatGPT do not satisfy their ‘authorship’ attribution criteria.

d. Responsible use. It is recommended that the use of these AI tools be limited to early stages of research, to inspire or suggest directions, not to produce content that will later be included in your deliverables. In any case, if you choose to transcribe texts produced by Generative AI, you must prove that you have rigorously checked such information against reliable sources, since Large Language Models (LLMs) such as ChatGPT tend to offer inaccurate, erroneous, and invented information.

4. Monitoring. The teacher may use tools to detect violations associated with the use of highrisk tools. Turinitin announced that it is working on adding an AI handwriting detection module. In addition, it is likely that, soon, the generative tools themselves will include a ‘watermark’ in their results that will make it easier to track whether a text was produced by AI. In any case, teacher will focus on pedagogical strategies rather than punitive ones.

5. Accompaniment of the teacher. At the beginning of the semester, the professor will explain what AI is, what language models are and how tools such as ChatGPT work, what opportunities and risks they generate for academic and professional work, and what ethical, environmental, and human rights implications are associated with the use of these tools. The professor will always be available to clarify the scope of these guidelines, to discuss and co-create them, and to resolve pointed questions about the use of AI.

4. Why is an informed, transparent, ethical, and responsible use of AI necessary?

- The use of AI tools must be informed, transparent, ethical, and responsible for –at least– four types of reasons:
 - (1) because AI tools are not always reliable;
 - (2) because there are risks that certain uses will negatively affect learning processes;
 - (3) because of the risk of users treating AI as if it were a human being; and,

(4) because the use of the tools has ethical and human rights implications due to the way they were developed and/or because some tools may replicate or amplify social issues such as discrimination. Each of the four reasons is explained in detail below:

- o First, because their answers are NOT always reliable even though, for example, an AI generator produces text that looks convincing. LLM-based systems such as ChatGPT do not perform with the accuracy of other tools used in learning environments such as calculators. In fact, ChatGPT tends to include false or fanciful information in its responses. Microsoft's built-in chatbot for its Internet browser, Bing, and demonstrations of Google's chatbot, Bard, have also presented the same types of problems. These systems do not distinguish true from false. Why does this happen? LLMs spin words from probabilistic inferences from the data they were trained on, but they do not have the ability to understand what they produce or associate meanings to the words they utter (they are 'stochastic parrots'). Recently, a media outlet that used a ChatGPT-type tool to write texts had to publish corrections to multiple articles due to serious inaccuracies.
- o Second, because the use of text-generating tools may discourage students' motivation to write and think on their own. It is worth reiterating that the learning activities in this subject seek to develop their cognitive skills and that this policy seeks to prevent some AI tools from becoming automated plagiarism mechanisms.
- o Thirdly, because of the risk that users, consciously or unconsciously, treat AI behavior as if it were human (Eliza Effect). LLM-based tools do not understand their output, but simply mimic language patterns from the synthesis of large volumes of data from which the program generates word sequences. This problem of anthropomorphizing machines is exacerbated by the fact that some chatbots have produced violent or harassing responses that could lead people to engage in harmful behavior.
- o Fourth, because of the ethical and human rights implications associated with the use of certain AI systems given that: some tools tend to reproduce or amplify derogatory and discriminatory stereotypes associated with gender, race, ethnicity or disability; technologies could have been developed from massive copyright infringement; new forms of colonialism through the non-consensual extraction of information from historically marginalized communities; some tools would have been developed in contexts of labor exploitation; the development and operation of such systems generates a considerable carbon footprint; and, the potential violation of privacy and personal data protection rights of those who use them.

5. Open-Guidelines:

- AI is a rapidly changing set of tools, which is why these guidelines will remain open to future evaluation, modification, and revision. From the beginning of the semester and throughout the semester the professor will open spaces to discuss the guidelines with the students and, if necessary, modifications to this guideline can be introduced through co-creation exercises.
- For those interested in the challenges that generative AI creates for those of us who teach in universities, I recommend this document by Professors Anna Mills and Lauren M. E. Goodlad.
- For those who would like to read a brief introduction on how LLMs work and what their main risks are, I recommend the article ‘On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ’ by Emily M. Bender, Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell.
- Anyone interested in accessing material to learn basic information about AI produced by MIT, I recommend checking out the DAILY Curriculum project.

32. Guidelines for the use of ChatGPT and text generative AI in Justice

I. Executive Summary

The emergence and general public availability of disruptive models such as ChatGPT, GPT-4, Bard, DALLE-2, Midjourney and Stable Diffusion, has given rise to a new era of creation and manipulation of text, image and videos. This is an evolutionary leap in artificial intelligence, based on the generation of synthetic digital content.¹ Generative AI is a subfield of research and development within the artificial intelligence ecosystem. It is evolving rapidly and focuses on generating images, music, text, videos, voices and computer code, from natural language text input provided by the user (prompt or instruction). The new AI assistants also make it possible to make multiple legal and judicial tasks more efficient, while at the same time they can enhance various aspects linked to the underlying legal reasoning. From the daily management of simple procedural acts, to the possible generation of documents (e.g. certificates, official letters, orders, notes, dispatches and resolutions, among others). Also, these large language models (LLM) such as ChatGPT, can address other more sophisticated tasks that carry multiple challenges and associated risks. The massification and making generative AI available for free for a vast public, has to be

qualified by the sensitivity of the state function performed by those who make up the Judiciary. Various constitutional principles and rights come into play here, while the principle of competition could be tense or violated. The approach that we propose in this guide is based on the following five postulates:

1) Literacy, awareness and permanent training based on the evolution of the state of the art of generative AI. 2) Responsible use, based on mitigating risks associated with large generative AI language models. 3) Human in the Loop: Essential human control before and after. Prohibition of delegation of decision making. 4) Understanding the intrinsic and external limitations of LLMs such as ChatGPT (inventions, hallucinations, inconsistencies, negative biases, among others). 5) Strategic impact assessment according to the scope of application.

In recent times, situations of inappropriate use have been made known, such as the case of the lawyers in the “Mata vs. Avianca Airlines., Inc.”, which went viral because the representatives of one of the parties had cited a series of non-existent judicial precedents to support a claim, resulting in them being required by Justice to give explanations in this regard. On that occasion, the aforementioned lawyers admitted to the court that it was an error due to the misuse of ChatGPT by one of them, a professional with a long career, who relied entirely and trusted the content generated by the system, without subsequently verifying that result.² These types of situations have paved the way for other judges, such as the case of a federal judge in the Northern District of Texas, to establish, as a rule, the need for the presentation of an explicit and specific statement by the lawyers, in which they indicate that no part of their presentation was written by an AI or, if so, inform that the content has been verified by a human being as to its accuracy and veracity of sources, prior to being presented to the court.³ This example shows that the incorporation of text generative AI as support for the execution of daily legal or judicial tasks brings with it a great challenge for organizations. Every new technology allows us to discover new kinds of uses while entailing new responsibilities.⁴ Generative AI is no exception. Based on the work we do at UBA IALAB on the impact of ChatGPT in different areas, including the legal field, we believe that it is a critical moment to carry out a “controlled and strictly supervised experimentation”, ushered by guidelines that are updated and modified over time with constant feedback from different stakeholders. The MIT Working Group is also working with this same line of thought in regards to the use of generative AI in the legal field.⁵ Below we take the first step to establish a general

guide, in the form of guidelines and recommendations, as a starting point in relation to the responsible use of this disruptive technology.

II. Goal

We propose a series of guidelines and recommendations for the responsible, ethical, appropriate and diligent use of text generative AI in the justice service. This type of technology is what is behind the systems that are available to the general public through different applications such as ChatGPT, GPT4, Bing and Bard, among others, some of which are free to access. Specifically, it is about making known the inherent limits and possibilities of these intelligent systems, and then providing guidelines so that users in the justice sector can use them in line with the duties and values that must be protected in the exercise of their function. Here you have to balance the possible benefits with the risks and potential harms. For the rest, this first version of the guide is developed under an open and iterative logic, in order to make it known and invite a broader and more diverse set of interested parties to participate in its eventual improvement and completeness, at the same time than to keep it permanently updated (See point X).

III. Conceptual framework.

Possibilities and limits of large language models LLMs emerged strongly towards the end of 2022 with the launch and general public availability of ChatGPT, and then GPT-4 in March 2023. These two models, from the company OpenAI, paved the way into a paradigm shift. On one hand, the provision to the user of a tool with the potential to transform various industries and sectors; on the other, democratization for access. Anyone with a device and an internet connection can now use them freely, regardless of whether or not they understand the underlying technology. LLMs, such as ChatGPT and GPT-4, are AI systems that can process natural language in the form of a conversation. Under this logic, they allow the user to generate textual content in various formats (e.g. poetry, news, questions, summaries, etc.) from an instruction or requirement provided to the system in natural language, commonly known as prompt. From a more technical perspective, they are generative AI systems⁷ trained on text string prediction tasks. Meaning they are trained to predict the probability of occurrence of a token (which can be a character, a word, or a string of words) given its preceding or surrounding context.⁸In the case of ChatGPT, it is a generative

model based on the Pre-Trained Generative Transformers (GPT) architecture, which can process sequences of elements (such as the words in a sentence) using a deep learning architecture that makes it easier for you to pay attention to the different parts of a sequence of words while processing it, achieving greater efficiency and precision in your tasks.⁹ ChatGPT has been fine-tuned using Human Feedback Reinforcement Learning (HFRRL) technique, whereby human annotators teach the model to be more accurate by rewarding and penalizing its results. This circumstance means that, from its training, the model is subject to the subjectivities of our species.¹⁰ As recognized by Sam Altman, CEO of OpenAI, and as we have verified from the IALAB in the research carried out this year¹¹, LLMs in general, and ChatGPT and GPT-4 in particular, are often imperfect and limited by design. That is, they are deployed to the world with knowledge of their defects so that society adapts to the product. In fact, within the limits of these systems, it is noted that:

- a. Sometimes they give coherent and convincing answers, which imitate confident style and expert jargon, but are incorrect or false (they hallucinate).
- b. In their responses they sometimes reflect prejudices, stereotypes, beliefs and negative social values present in their training data (bias).
- c. Sometimes they are not solid and fail relatively frequently when given tasks that involve logical reasoning.
- d. These are systems that are very sensitive to adjustments in the formulation of input phrases or sentences.
- e. Some of these systems have knowledge limited to a certain date (e.g. ChatGPT).
- f. They are excessively detailed and over-explain.
- g. They make assumptions about facts.

Arthur C. Clarke many decades ago stated: “Any sufficiently advanced technology is indistinguishable from magic.” In the judicial branch, this statement is not sustainable from any point of view when evaluating the usefulness of a tool or technology. In fact, they are not artificial oracles who know everything and can answer everything well. Also, another fundamental starting point to consider is that these LLMs deepen the notion of “black boxes” that we have worked on in multiple investigations, projects and conferences. This trait has two relevant dimensions to consider. On one hand, it is not known with what data the system was trained. There are various hypotheses about the sources used, but there is no reliable information that specifically indicates which databases they use to generate the answers. In this dimension, we also do not know what parameters the company applies to the language models. That is, the rules that adjust the outputs of a system. An obvious example might be this: if I ask it to tell me how to be Hitler, the system

refuses to answer. This is not “natural” in the system but is a specific configuration made by the programmers. In another dimension, the black box responds to the way in which data is processed. It is a subspecies of artificial neural networks that mathematicise knowledge based on probabilistic criteria. While they process the data, they do not know the specific step by step of how they arrive at a certain result. We have explained this phenomenon in various articles and research-.¹² This ultimate characteristic makes it difficult and, in certain cases, eliminates the possibility of explaining with adequate detail the reason for the decision, in order to justify and properly motivate decisions in cases where this is constitutionally or legally required. Ultimately, this prohibits any alternative to decisional delegation and, in some way, limits the role of this type of tools to that of an assistant

IV. Possible uses of text generative AI in Justice

The use of text generative AI in Justice can be presented as a genre that includes two species: 1. Transversal uses: they can occur in relation to any procedural stage and instance and in relation to any writing; 2. Specific uses: they are presented for specific tasks or documents that must be planned and dictated in the different procedural stages, instances and jurisdictions (specialty criteria of the subject involved). Below we will provide a list of possible transversal uses of AI in Justice, as well as specific uses. This list, although broad, is merely illustrative, given that this subfield of generative AI is dynamic and is constantly developing and evolving. It is key to promote continuous improvement to reveal other possible uses, just as it is important to consider those that should be discarded, and it is also critical to make constant reviews about precision, coherence, hallucinations and all issues linked to the limits of large models. of generative AI language.

1. What would be some possible transversal uses of generative AI in Justice?

Transversal uses are those that can occur independently of the procedural stage and instance that a judicial case goes through and in relation to any writing. It includes the pre-judicial stage as that occurs when the process is ruled. Among the possible transversal uses identified, so far, we have found the following: a. Search for existing information (National Constitution, treaties, doctrine, jurisprudence, etc.). b. Search within texts (National Constitution, treaties, doctrine, laws, jurisprudence, etc.). c. Request for ideas or alternatives to problems or judicial conflicts.d.

Synthesis of different types of judicial documents (e.g. lawsuits, pleadings, sentences, doctrine, etc.). e. Interpretation, evaluation or weighing of legal rules or principles applicable to a specific case. f. Generation of models or templates for resolutions (e.g. simple rulings, opinions, etc.). g. Make analogies or metaphors about arguments, claims or possible explanations linked to the content of a judicial decision. h. Combination of legal information with information from other disciplines (e.g. art, literature, cinema, etc.). i. Conceptual distinctions and combination of legal arguments. j. Strengthening and deepening of legal arguments that are presented as a starting point for the system. k. Analysis of judicial or legal documents (e.g. contracts, writings, demands, resolutions, etc.). l. Compare data or information between resolutions, legal regimes or other documents. m. Assessment, interpretation and/or consideration of the origin of certain claims. n. Respond to emails, notes and letters. o. Translation of documents. p. Improve writing, apply clear language/inclusive language/synthesize/remove gerunds.

2. Which would be some possible specific uses to generative AI in Justice?

Specific uses are those that are presented for specific tasks or writings that must be planned and dictated in the different procedural stages, instances and jurisdictions (specialty criteria of the subject involved). Among the possible specific uses identified, so far, we have found the following:

- a. Redaction of drafts of simple orders and sentences.
- b. Relationship between legal texts provided by the user (e.g. demand and response to the demand).
- c. Relationship between judicial texts and texts provided by the user (e.g. sentence and appeal).
- d. Identification of the claims and requests made in writings (e.g. in a complaint or in a response).
- e. Enumeration of the means of proof proposed in writings.
- f. Suggestion of new means of proof or expansion of means of proof.
- g. Summary of the object of the claim.
- h. Suggestion for improvements in the writing and content of judicial texts.
- i. Analysis and evaluation of evidence produced in relation to the facts under debate and claims of the parties

V. Relevant considerations regarding the potential uses of text generative AI

It is important to keep in mind that the identification of possible transversal and specific uses carried out in this guide does not necessarily mean that its implementation will yield correct, precise, safe, useful and/or adequate results. Indeed, the current state of evolution of text generative AI, initially addressed in the book “ChatGPT vs. GPT-4: imperfect by design?”, shows that we

still face systems with important limitations that many times, to achieve good performance, require a user with good prompting skills. That is, with knowledge and skills that allow them, through the instructions they provide to the system, to guide it in carrying out the task of processing and generating natural language, with or without legal content, to obtain more efficient, useful, effective and accurate results. Notwithstanding this, it should not be lost of sight that sometimes the limits go far beyond the capacity of the human person to interact with the system.

VI. Specific guidelines and recommendations for the use of AI in Justice

VI.1. Introduction

The responsible, ethical, appropriate and diligent use of text generative AI by judicial officers demands a holistic approach that integrates specialized knowledge; transparency; human supervision; regulatory compliance; protection of data and information of the organization and third parties; addressing challenges related to biases and maintaining constant and ongoing critical and ethical judgment on the part of the system user. In essence, a balanced approach must be adopted between the adoption of generative AI, which is presented as a work tool with the capacity to enhance, amplify and complement human intelligence, and the preservation of the fundamental duties, values and principles that go through the judicial function. Firstly, these guidelines have as their conceptual framework the map of ethical documents that has been synthesized and that form part of the book titled “Tratado de Inteligencia Artificial y Derecho”, published by Thomson Reuters-La ley, in November of 2023.¹³ From this perspective, we propose an illustrative list of guidelines that every judicial agent, regardless of their position or hierarchy, should comply with to make responsible, ethical, appropriate and diligent use of text generative AI when using it as a tool at work. At the same time, others that are stated aim to promote that people who use generative text AI in Justice can make the most of the capacity of these systems and obtain more accurate, useful, precise and satisfactory responses and, at the same time, optimize the interaction with the AI system to make it more efficient (fewer interactions/better results). The latter ones will help agents make ChatGPT or other AI that are used, a useful, agile and versatile assistant for the development of some of your daily tasks that involve writing text in different formats, analyzing and solving problems and synthesizing documents, among others. In all cases, the proposed guidelines will allow agents to make ChatGPT or other AI that are used, a useful assistant for the

development of some of your daily tasks that involve writing text in different formats, analyzing and solving problems, and synthesizing documents, among others. Below, the identified guidelines are specified and grouped. Also, we provide, as a suggestion, some specific recommendations that can be adopted by organizations to make them effective in their daily practice

VI.2. List of guidelines and recommendations

i. In relation to the data of the organization and third parties

i.1. When using documents or information of the organization as input: protect and ensure confidentiality

Court documents often contain sensitive, confidential or private information, so we must ensure that we implement appropriate security measures to protect the privacy and confidentiality of the organizations and third-party data contained in those documents while using generative AI. The use of text-generative artificial intelligence in the judicial field should not lead to the processing by third parties of personal data, nor organizational data, so it is important to comply with data protection regulations, but also with the rules of reserve of the judicial power itself to guarantee the confidentiality of the information. Implementation Tips:

- Design, together with specialists in data protection and information security, a scheme of best practices in risk management to guarantee the confidentiality of the information of the organization and third parties contained in documents and systems.
- Best practices should include techniques for users to learn how to implement data anonymization mechanisms to ensure the protection of the organization and third-party information.
- Design awareness campaigns and ongoing training for users on the security and privacy risks associated with text-generative artificial intelligence used in the judicial field.
- Establish strong confidentiality agreements with agents to safeguard the data they manage.

ii. In relation to the people of the organization

ii.1. Analyze the level of knowledge of users and design awareness and knowledge plans for generative AI

The awareness and knowledge of text generative AI among the people who are part of the organization are essential to harness its potential; ensure appropriate, ethical, legal and successful use of this tool, and avoid overconfidence. It is important that the organization's authorities know the degree of knowledge that the agents have about the operation, possibilities and limits of text or image generative AI. A holistic view of the level of knowledge of the organization will allow you to identify training needs and appropriately address existing knowledge and skill gaps. This

can be done through the use of various strategies and means of communication, such as campaigns, talks and training that focus on explaining the three verticals referred to (functioning, possibilities and limits), as well as the human roles and responsibilities in the process, the risk prevention and best practices. Implementation Tips: - Design and implement specific communication campaigns on text generative AI in the Judiciary. - Implement training and continuing education programs for different levels of users (initial, intermediate, advanced) focused on understanding technology and its ethical implications. - Provide a clear description of the guidelines and recommendations for the use of text generative AI in the Judiciary. - Train on risk management policies and practices adapted to text generative AI. - Emphasize the role of people in validating and verifying AI-generated texts. - Establish multidisciplinary consultation teams so that they can address the study and work of complex use cases.

iii. About the results of the AI system

iii.1. Iterate responses with appropriate human supervision focused on manually reviewing and editing responses

Human supervision is essential to ensure the accuracy, quality and ethics of AI-generated responses, correct and improve them, prevent bias, inappropriate, incorrect or illegitimate content. Therefore, it is important to permanently emphasize that human beings will always continue to be control and decision-making agents. This will mitigate fears and increase confidence in the technology, while collaborating in its legal, ethical, responsible and appropriate use. Systems such as ChatGPT, GPT-4 and Bard, among others, have been designed specifically for text generation tasks in conversational format. They are not optimized to be used as a search engine. These are models with an adequate structure to generate coherent answers to questions and hold natural conversations, so they may present limitations if they are used to search and retrieve specific information with a precise and reliable result. This is why the use of text generative AI, at least for the moment, can lead to the generation of coherent and convincing answers, which appropriately use technical language, but which may eventually be inaccurate, incomplete or based on outdated or false information. Implementation Tips: Ensure adequate human supervision to: - Always use verification mechanisms and subsequent human validation of responses to guarantee their accuracy and reliability. - Always use verification mechanisms and subsequent human validation of responses to control the sources cited when dealing with systems connected to the Internet. - Correct errors, improve quality and ensure the appropriateness of responses. - Check that they do

not contain false information. - Check that they do not contain inappropriate or offensive content. - Ensure their ethical adequacy (e.g. checking that they do not reflect negative stereotypes or prejudices, etc.). - Ensure its adaptation to current regulations and to the specific case according to the facts. - Improve your precision. - Approve methodologies for the preparation of prompts appropriate for each case.

iii.2. Consider AI-generated documents as drafts

In addition to the limits indicated, related to the risks that these systems present in terms of generating erroneous, biased or meaningless responses, it must be kept in mind that LLMs act as "black boxes," which brings serious difficulties. In principle, it is not possible to understand and justify how the AI arrived at a specific result or decision. As a result, when they are used to generate judicial documents, they should always be used as an assistant or co-pilot to generate a draft that will speed up the task of final preparation of the document which, in all cases, must be left to a human person who will control its content by applying its expertise. Implementation suggestions:

1. Define specific guidelines for the use of generative AI in the drafting of judicial documents (e.g. in which cases, with what prompts).
2. Work on the design of appropriate prompts, that provide the system with a context, provide terminology, contextualize situations and provide examples, among others, to improve the quality of the drafts generated.
3. All drafts generated by AI must be reviewed and edited by people qualified in the subject matter being worked on. This is to ensure their accuracy and consistency with applicable standards and rules.
4. Regularly monitor and evaluate the performance of generative AI used in the generation of judicial documents through user feedback in order to introduce improvements in the way it is used and identify new use cases.

iii.3. Evaluate and mitigate biases in the responses generated

Text generative AI systems can produce biased responses that reflect gender stereotypes, prejudices, beliefs, or negative social values,¹⁴ so people who work in the Judiciary must be aware of this problem and have tools to evaluate and eliminate them from the results generated, in order to mitigate their negative impact, avoid discrimination or unfair differences in treatment and representation, and ensure fairness in the use of generative AI. Implementation suggestions:

1. Train and sensitize agents of the Judicial branch so that they understand how generative text AI systems work and the problems they present in relation to bias.
2. Encourage collaboration between experts in AI ethics and judicial ethics to develop protocols for the evaluation and mitigation of bias in AI-generated responses.
3. At the individual level, carry out a rigorous evaluation of the

responses based on the applicable ethical and regulatory standards to ensure they are impartial and do not incur in discrimination or unfair or unacceptable differences in treatment and representation of groups. 4. At the team level, ensure careful review of AI-generated responses, especially when used in the context of the exercise of judicial function, to corroborate the impartiality/objectivity of the generated response and correct any bias or inaccuracy identified.

iv. Regarding prompting

iv.1. For prompting in general In the context of text generative AI, a prompt is an instruction or initial phrase that is provided to the LLM to generate text and achieve specific results such as writing texts in different formats (e.g. news, emails, poems, etc.), automatic translation, dialogue generation, among many others. In other words, it is a text input provided in natural language by the user of the system that serves to guide the latter in generating the desired response. For example: "Answered the email that I am going to provide you below in a gentle but formal tone, in a maximum of 10 lines. Informal to the person that I will not be able to attend the meeting on 7/1/2023 because I have other commitments." The prompt is a way of communicating with text generative AI systems, so it is important to adjust and formulate them appropriately to obtain better results. Also, it is important to be aware that each model can respond differently to the same prompt, and that the same system can generate new results for the same prompt. Specific suggestions:

- Be clear, precise and concise: provide clear, specific and, to the extent possible, concise instructions. Example: Unclear and unspecific prompt: "Write me 15 lines about AI and Law."/ Clear and specific prompt: "Write me 15 lines about AI and Law. In particular about the relationship that exists between both disciplines, what are the points of contact between both and how they influence each other."
- Provide sufficient details: provide the details so that the system can better process the instruction and more appropriately comply with it. Example: Prompt not detailed: "Write me 15 lines about music." / Detailed prompt: "Write me 15 lines about the influence of national rock on Argentine culture in the 1980s."
- Set a context: provide an adequate context so that the system can adjust to the topic or situation raised. Example: Prompt without context: "Write an email to invite to a meeting." / Contextualized prompt: "Write an email in a gentle and formal, brief tone, to invite Juan Pérez's work team, to his office, to address different topics that have been worked on in the last month."
- Indicate recipients: indicate who are the intended/target users of the text to be generated. Example: Prompt without intended/target users: "Summarize this text using clear language." /Prompt with intended/target users: "Summarize this

text using clear language. Aim it at 15-year-old teenagers.” They could also be: non-professionals; kids; teenagers; older adults; etc. - Set a tone: indicate the tone that the text to be produced should have. For example: formal, informal, funny, persuasive, assertive, etc. - Use examples and keywords: include examples or key words to clarify and illustrate the objective sought with the instruction. Example: Prompt without examples: “Describe a scene of two children playing.” /Prompt with examples: “Describe a scene of two children playing. For example, how the brothers from the series “Loud” play House.” - Experiment and iterate the prompt: sometimes, minimal adjustments in the instruction or requests for clarification, precision and provision of new details to the generated result, through the use of new prompts or the iteration of the one you are using, lead to more precise and interesting results. - Indicate a role: request the system to assume a specific role prior to giving the instruction. Example: Prompt without role: “Describe a scene of two children playing.” /Prompt with role: “Assume the role of a science fiction writer for children. Describe a scene of two children playing. For example, how the brothers in the series play Loud House.” - Request a format: indicate the desired output format for the text. For example: draft, writing, outline, comparative table, dialogue, etc. - Indicate an output writing style: request the desired writing style for the text to be generated. For example: clear language, inclusive language, inclusion of gender perspective. Example: Prompt without asking for style: “Describe a scene of two children playing” / Prompt with style: “Describe a scene of two children playing, apply a gender perspective when considering the game they play.” - Citation of sources: for systems connected to the internet (e.g. Bing), request citation of all sources that support the content of the generated response and then check them.

iv.2. For legal prompting

A legal prompt is an instruction or initial phrase that is provided to the LLM for the generation of text related to legal matters. The legal prompt can be especially useful for applying generative AI to tasks related to the legal field such as drafting contracts, analyzing facts, generating legal arguments and counterarguments, among others. Suggestions specific to achieve a good legal prompt:¹⁵ - Be specific and clear: Provide precise and clear instructions on the legal topic you are going to work on. - Set the context: Include relevant information to provide adequate context to the system, such as, for example, details of facts to be analyzed, non-relevant circumstances, applicable principles, legal references, among others. - Assign a tone: Specify the desired output tone for the general text, appropriate to the task being performed. For example: formal, informal,

assertive, persuasive, funny. - Request a format: Request the desired output format for the text. For example: draft, writing, outline, comparative table, dialogue, etc. Indicate a role: Request the system to assume a specific role (“act as”) prior to giving the instruction. For example: role of judge, role of specialist in law and technology, etc. - Provide the desired objective: Indicate to the system what goal or purpose is sought to be achieved with the response. For example: inform, reconcile, resolve, analyze, etc. - Include keywords: Incorporate keywords into the instruction to direct the response generated towards the desired objective. For example: appeal, demand for support, arguments in a labor trial, oral arguments, etc. - Set limits: Specify restrictions on the answer. For example: number of words or characters, line limit, do not use generic masculine or grammatical masculine. - Indicate the target audience: Define who the target recipients of the text are to achieve personalized content. For example: “write a summons for an interview to make contact with a minor.” - Request fonts: In models connected to the Internet (e.g. Bing), request the citation of all sources that support the content of the generated response and then check them. - Request consideration of multiple perspectives and counterarguments: Ask you to take on more than one role and consider multiple points of view, opinions, and counterarguments on the same topic. See example of superprompt technique. Example of a prompt for compliance: Hello ChatGPT, [insert text describing the facts of the case in sufficient detail].I need you to determine the legality of [insert the text that describes the facts to be analyzed in sufficient detail], based on [indicate the regulations under which the analysis is intended to be carried out]. It is likely that after the first response we must iterate with the AI to refine and/or adjust the result, as well as to clarify doubts or unresolved questions that arise from that initial response.

iv.3. Legal prompting for procedural acts

Text generative AI can also be used as an assistant for dictation of procedural acts that respond to requests made by the parties during the process. This allows not only to simplify the way in which the response is submitted to an office, but also to reduce its length and also to create models that can incorporate numerous determining procedural variables. To obtain accurate, controlled, coherent and reviewable results it is advisable to have a conversation which complies with the specific recommendations for a good prompting legal, applied to a conversation context using the methodology suggested below. Specific suggestions to achieve a good legal prompt for procedural acts: Prompt to generate simplified order models: Hello ChatGPT. Simplify the following text [office model], maintain a wording that allows it to be used as part of a model, use legal language,

keep data that refers to telephone numbers, law numbers, deadlines, amounts of money, names of entities. Prompt to generate office models with options: 1. Because generative agents do not have specific procedural knowledge, and because they can respond to conceptual explanations, dispatch models can be established with indications of the procedural circumstances that would motivate its dictation. In this way, you can approach a conversation with the system that will allow you to arrive at a useful response in the course of the conversation. A first prompt model with numerous determining procedural variables is provided below: Hello ChatGPT. I am going to assign you two prompts to work with options later depending on the next step that the process must follow. Option 1: [“in case of compliance with the presentation of... it is resolved...”]. Option 2: [“in case of not complying with the presentation of... it is resolved.”]. 2. Taking as a starting point a prompt with numerous procedural variables that determine the dictation of one or another decision, the system can be asked to prepare a list of questions that constitute the procedural reasoning on which a decision must be made. The result is more precise if the number of options is indicated. A second prompt model is provided below: Based on the following text [prompt with numerous determining procedural variables], prepare a questionnaire of the questions that must be taken into account in each of the XX options. 3. If the determining procedural variables are appropriately written and the questions in the options questionnaire are properly answered, an automatic response from the AI can be achieved, consistent with the presentation of the party and the procedural moment being passed. To motivate a response of these characteristics, it is necessary to request it together with the answers to the previous questionnaire. A third prompt model is provided below: Now I need you to use the answers from the following list [responses to the procedural variables questionnaire] and complete the dispatch model [dispatch/providence model with determining procedural variables]. 4. It is possible that the result of the automated base model is procedurally coherent, but it is also possible that it lacks the usual writing modality of the person who signs the order. To resolve this obstacle, the model can be asked to imitate the writing style of other offices and adapt the generated model to achieve similarity. This correction can be considered the final step of the automation process for generating draft dispatches, so it is the right time to carry out all the corrections and supervisions inherent to the professional work. Below the model prompt style appropriateness: Based on this first text [decree/dispatch/providence model with the style of the person signing the dispatches] rewrite the second text [automated base decree/dispatch/providence model] imitating the way of writing of the first text.”

v. For the approach, analysis and resolution of legal and non-legal problems

v.1. Super Prompting for exploring diverse approaches to a problem

It is a technique of prompting that matches multidisciplinary logic with a response approach based on Tree of Thoughts. It is done through a super prompt that contains several commands or indications together that optimize the dialogue with the LLM. To do this, it proposes the intervention and response, in a cascade and jointly, of more than one expert in one or several topics (there may be two or more iterations). This interaction methodology allows improving the exploration of coherent text units (thoughts) that serve as intermediate steps for problem solving; optimizes the human/AI dialogue, since it allows generating a more complete response (the AI assumes three roles at the same time and confronts them in deliberation); expands the possibilities of thinking about a phenomenon based on the “simulated dialogue” between the predictions and the generation of synthetic data behind each “expert”; exponentially increases productivity in the face of complex tasks such as the development and analysis of process maps, task maps, decision trees and conceptual matrices based on various perspectives; optimizes the probabilistic approach, the representation and simulation of a part of human knowledge provided by generative AI. Tips for using the super prompt:

- Clarify and segment orders in the same super prompt
- Respect punctuation rules
- In the first prompt include all the orders and anticipate that the topic will be given in a second prompt
- In a second prompt introduce the theme
- Use the “conceptual matrix” formula
- During the interaction, if the model releases the fulfillment of the proposed task, indicate “ok, continue.”

Below is a basic model of super prompt prepared from UBA IALAB:

Hello ChatGPT. I need you to act as three experts in ___, ___ and ___. They must debate the problem or issue that I will indicate to you. The debate must be iterated and I ask you to reflect the arguments in a table. I need you to then reach an agreed conclusion with a proposal or recommendation and to develop a conceptual matrix. Then I'm going to ask three other experts in ___, ___ and ___ to put the arguments and conclusion into crisis. Below I write the problem or question...

v.2. Strategies for evaluating evidence and seeking suggestions for resolving cases

Text generative AI can also be used as an assistant for evaluating evidence and facts to obtain suggestions or alternatives for resolving cases; for help to detect inconsistencies or contradictions in the evidence and facts presented; to evaluate argumentative coherence and identify inconsistencies in the arguments presented; to identify key points of the facts of a case and draw

conclusions and also to identify and generate counter arguments or alternative points of view in order to consider different perspectives on the topic under analysis. Specific suggestions for evaluating evidence and possible determination of the resolution of a case 1. First of all, there is a prompt through which the model is informed that we request its help to make a decision in a complex case. We outline a brief summary of the case and indicate that in a subsequent prompt we will inform of the evidence and facts. The purpose of the provided texts is to obtain the help of ChatGPT for the evaluation of evidence and facts. a. A first model is provided below prompt: Hi GPT-4. I hope you are fine. I need your help. I am a judge [insert legal specialty] and I have to make a decision in a complex case that I have to deal with in the court that I am in charge of. The case [insert text summarily describing the facts of the case]. I will discuss the evidence and facts in more detail in the next prompt. Could you help me evaluate the evidence and facts so that I can determine [insert text explaining the intended objective. For example, if the dismissal was due to a cause attributable to the plaintiff or was it due to a cause attributable to the defendant who fired her due to the pregnancy]? b. Then we proceed to provide the model with a second prompt which details the facts and evidence. [insert text that describes the events in sufficient detail]. As a judge I must evaluate these facts and evidence to decide [indicate what decision you should make to provide context. For example, if the dismissal was for a cause attributable to the plaintiff or it was for a cause attributable to the defendant who dismissed her due to the pregnancy]. Can you help me? 2. Prompt model to carry out an analysis of coherence and consistency between evidentiary means. For example, between testimonies. Hello ChatGPT. I ask you to review and evaluate the logical and argumentative coherence between the testimonies A, B and C that I am going to provide you. Identify points of agreement and points of contradiction between them. I ask you to put them in a box. Below I give you the testimonies: [insert the text with the testimonies of A, B and C].

vi. To work on legal and non-legal texts

vi.1. Text simplification

Text generative artificial intelligence can be very useful to simplify technical language and generate a text expressed in more accessible terms (clear language) and/or free of masculinities (inclusive language), with the aim of improving and facilitating communication with the citizens. Models of prompt to simplify a text and propose an output in clear and/or inclusive language: 1. I need to simplify the following paragraph. Use legal language, keep data that refers to law numbers, deadlines, amounts of money, names of entities data or people: [insert the text you want to

simplify]. 2. I need you to explain to me the text that I am going to provide you below, and that you do so using clear and accessible language for a non-specialized audience. I ask you to maintain the formal tone, preserve the legal language as much as possible and bring warmth to the writing: [insert the text you want to simplify]. 3. I need you to explain to me the text that I am going to provide you below, and that you do so using clear and accessible language for a non-specialized audience: [insert the text you want to simplify]. 4. I need you to write this text eliminating masculinities, while also using clear and accessible language for a non-specialized audience. I ask you to maintain the formal tone, preserve the legal language as much as possible and bring warmth to the writing: [insert the text you want to simplify].

vi.2. Text summary

Text generative artificial intelligence is useful for summarizing legal and non-legal texts, because it allows this task to be carried out with good precision and very quickly. In this way, the model can be required to identify and extract relevant information, find keywords, and discard unimportant content. Two models of prompts are provided below to summarize a non-legal and legal text: 1. Hello ChatGPT. I need you to summarize the key points of the text that I am going to provide below. You must also provide me with the conclusions: [insert the text you want to summarize]. 2. Hello ChatGPT. I need you to summarize the key points of the text that I am going to provide below. I also ask that you identify the relevant facts and legal arguments, and provide me with conclusions: [insert the legal text you want to summarize].

VII. Team:

General Management: Juan Gustavo Corvalán and Mariana Sánchez Caparrós UBA-IALAB research team: Giselle Helleg, Carina Papini, Melisa Raban and Antonella Stringhini. Dyntec-Lab research team: Marco Rossi, Franco Orellana.

VIII. Do you want to be part of the Guide?

Scan the QR and leave us your contributions From UBA-IALAB we want you to be part of the construction of this Guide. To do this, we enable this participatory, open and transparent space with which we begin the public consultation so that you can make your suggestions and propose

new uses, models of prompts, guidelines and recommendations not included in this first version, that you understand are relevant to meeting the objectives of the document.

33. Guidelines for the use of artificial intelligence in university contexts

Introduction

- We use artificial intelligence tools daily: Consciously or unconsciously, practically all of us use some kind of artificial intelligence (AI) in our daily lives, in particular systems based on language models. For example, we use language models when we are writing a text message or an email and the “autocomplete” tool suggests how to finish writing the sentence. We also use language models to support us in carrying out academic activities, such as automated translation systems or programs that detect grammatical errors, among others. In addition, nowadays, there are web platforms and mobile applications that provide access to AI tools to generate texts and audiovisual content
- Some uses of AI tools can support learning, research, teaching, and other professional activities: There are appropriate uses of such technologies, which can not only be useful to support learning, teaching, and research processes at the university but also at a professional level. For example, different AI tools nowadays contribute to activities at all stages of public policy processes, and, in Colombia, more than 50 public sector entities have adopted automated decision systems to support the performance of their functions.
- But certain uses of AI are risky, so their use must be informed, transparent, ethical, and responsible: AI tools are not suitable for every type of activity and certain types of use can be counterproductive to the pedagogical process. Some uses of AI can generate risks for users and third parties. Furthermore, as explained in section 4 of these guidelines, it is essential to be aware of the ethical, environmental, and human rights implications associated with the use of these tools. It is precisely because of these risks that some universities and professors have published guidelines, recommendations, principles, policies, and directives on the use of AI in academic contexts (see list in section 5 of these guidelines).
- We should not assume that all our students are “digital natives” or that they want to adopt new technologies: . People's attitudes towards technologies are diverse and complex, it should not be assumed that all our students are eager to adopt new technologies. For example, in the surveys I

have carried out in workshops with students on the use of AI at university, I have found that these systems generate very diverse feelings in them: fear, intrigue, excitement, uncertainty, curiosity, distrust, euphoria, worry, amazement, expectation, anguish, tranquility, confusion, fascination, caution, questioning, rejection, uneasiness and comfort. Moreover, people's capacity to adopt new technologies is heterogeneous and depends on different economic, social and cultural factors, among others. All the above implies that those of us who teach should implement differentiated teaching strategies according to the needs of our students and that we should not force the use of new AI systems.

- Guidance for the use of AI inside and outside the university classroom: These guidelines provide basic guidance for university professors and students on how to use AI, particularly large-scale language models, in an informed, transparent, ethical, and responsible manner. The challenges of using AI in education are not limited to aspects associated with learning processes and academic research. For example, the adoption of AI to (semi)automate personnel hiring and evaluation processes or for the surveillance of educational facilities also has legal and ethical implications. But these issues are beyond the scope of these guidelines since the rules and principles outlined here focus exclusively on the use of AI for educational and academic purposes inside and outside the university classroom.
- Definitions of three key terms: The following are working definitions of three technological terms used in these guidelines: (A) Artificial Intelligence (AI) System: “a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions that influence real or virtual environments. AI systems are designed to operate with varying levels of autonomy”. (B) Language model: algorithms that process, analyze and produce natural language in written or spoken form. (C) Large Language Model (LLM): algorithms trained on a vast amount of information whose main function is to predict the most likely string of words given the context that precedes or follows it.

Objectives of these guidelines

- Contribute to digital literacy. Contribute to learning basic knowledge about the use of AI in university contexts – with emphasis on LLMs. This objective includes teaching about its benefits and risks, its implications for society, and the development of critical digital competencies.
- Promotion of responsible use of AI. Promote the responsible and ethical use of these technologies in learning processes and for future use in professional life.
- Prevent the violation of academic integrity. Prevent situations in which students consciously or unconsciously incur in academic dishonesty.

Rules for the use of AI in and out of the classroom

- General rule. The use of AI as a support tool to carry out different learning activities is allowed. The parameters for the use of these tools described below distinguish between “low risk” and “high risk” tools, depending on the risks that the respective AI generates for the pedagogical process, for users and for third parties.
- Rules for the use of low-risk AI. Students may freely use low-risk AI tools, i.e., those that allow to correct or review student-generated content or those that allow to collect and process data. For example, grammar correction tools, translation tools, audio-to-text transcription tools, and Internet information search tools, among others. Where such tools enable students to do meaningful work (e.g., translation), they must indicate their use in the appropriate section (e.g., methodology section). If students have doubts about how to classify the risk of using a specific AI, I suggest contacting their teacher.
- Rules for the use of high-risk AI. AI tools that are used to generate synthetic text are high-risk. That is, LLMs that can be accessed through chatbots, such as ChatGPT, Bard, Bing, and Claude, etc. Students can include AI-generated content in their individual and collective deliverables when four requirements are met: i. Informed use. Prior to using the tool, students must investigate who or what company developed the tool, how it was developed, how it works, what functions it can perform, what its terms of use are, what the company does with the information it collects from the user, and what limitations and/or risks it may generate for the user and/or third parties. ii. Transparent use. Students must explain in detail and expressly which tool were used and how the tools were used in the methodology section or in the section that they deem pertinent. This is a requirement currently requested by scientific journals and publishing houses, such as Nature and

Cambridge University Press, to whom manuscripts are submitted for peer review. iii. Ethical use. In relation to the use of LLMs, students must distinguish what was written or produced directly by them and what was generated by an AI tool. On this point, general citation rules apply, e.g., use quotation marks if textual paragraphs are included. For citing synthetic text produced by LLMs, students may consult the rules suggested by the APA style team. Violations of this rule, particularly with respect to the second requirement, will be dealt with under an approach like that which applies in cases of plagiarism. In addition, ChatGPT is not a person, so it cannot be considered a co-author. Some scientific journals have already updated their publication policies to clarify that language models such as ChatGPT do not satisfy their “authorship” attribution criteria. On the other hand, the ethical use of AI systems – in general, not just LLMs – should be informed by four principles of bioethics: beneficence (achieving benefits), respect for the autonomy of persons (including obtaining consent if third parties are involved), justice (especially nondiscrimination), and non-maleficence (not intentionally or recklessly causing harm to others). iv. Responsible use. The use of AI tools should be limited to early stages of research, to inspire or suggest directions, not to produce content that will later be included in students’ deliverables. In any case, if students choose to transcribe texts produced by an LLM they must prove that the content was rigorously checked with reliable sources, since these systems tend to offer inaccurate, erroneous and false information. Responsible use also implies avoiding that the use of the system does not cause damage to the user or third parties. For example, it is not recommended to include personal data or confidential information about the students or others in chatbot queries because once such information is entered, the user loses control of who can access it and there is a risk that it will be known by unauthorized third parties.

- Plausible uses of LLMs: i. Explore. To explore new issues and seek inspiration, i.e., use the chatbot in the earliest stages of research. ii. Automate basic tasks. To support exercises that involve recognition and/or reproduction of patterns in texts, for example: translating, summarizing, modifying the tone and style of the text – for example, to make the text clearer or more accessible to certain audiences –, correcting spelling and grammar, audio-text conversion, programming, etc. Students should no rely on automated tasks, the LLM's output must be carefully reviewed. iii. Trivial writing. To support unimportant writing tasks. To decide whether something is trivial or not, think about the following heuristic: “what would happen if my interlocutor found out that I used an LLM to write the message?”. iv. Sparring. Testing one's own arguments or ideas through

interaction with the chatbot, for example, by asking the system to respond with criticisms or counterarguments to the student's statements.

- Not recommended uses of LLMs: i. Use the chatbot as the main or only source to search for factual or technical information. I do not recommend it, LLMs are not reliable search tools. Besides, Internet search engines are more efficient and reliable for that task. In any case, since searching chatbots can be perceived as an easier and faster way to obtain information, I recommend that the information obtained through chatbots should always be carefully checked with reliable sources. ii. Use of synthetic text for substantive issues. I do not recommend it because of the tendency of LLMs to produce answers with inaccurate, erroneous and false information. In any case, if students do so, each point should be checked against other reliable sources. iii. Doing mathematical calculations. LLMs are not trained to perform mathematical calculations and thus often fail to answer simple arithmetic operations.
- Accompaniment of the teacher and development of digital competencies. The teacher will dedicate class time to explain what AI is, what are LLMs that power tools like ChatGPT, what opportunities and risks do these systems generate for academic and professional work, and what are the ethical, environmental, and human rights implications associated with the use of these tools. The professor will be available to clarify the scope of these guidelines, to discuss and co-create them, and to resolve specific questions about the use of AI. In addition, the teacher will seek to carry out activities that contribute to the digital literacy of students, particularly for the development and strengthening of digital competencies such as the ability to critically evaluate the benefits and risks of new technologies.
- Supervision. The teacher will use different sources to identify potential situations of academic dishonesty using AI, but he or she will not rely on the tools available to detect synthetic text since none of these tools are currently reliable. In any case to avoid risks associated with academic integrity, the main strategy the teacher will use will be pedagogical not punitive.

Why the informed, transparent, ethical, and responsible use of AI in academia is necessary?

- The use of AI tools must be informed, transparent, ethical and responsible for – at least – four types of reasons: (1) because LLMs are not always reliable; (2) because there are risks that certain uses may negatively affect learning processes; (3) because of the risk of users treating AI as if it

were a human being; and (4) because the use of the tools has ethical and human rights implications due to the way they were developed and/or because some tools may replicate or amplify social issues such as discrimination. Each of the four reasons is explained in detail below:

- First, because their answers are NOT always reliable even though, for example, an AI generator produces text that looks convincing. LLM-based systems such as ChatGPT do not perform with the accuracy of other tools used in learning environments such as calculators. In fact, ChatGPT tends to include false or fanciful information in its responses. Microsoft's built-in chatbot for its Internet browser, Bing, and Google's chatbot, Bard, have also presented the same types of problems. These systems do not distinguish true from false. Why does this happen? LLMs spin words from probabilistic inferences from the data they were trained on, but they do not have the ability to understand what they produce, nor do they associate meanings to the words they output (they are “stochastic parrots”).
- Second, because the use of text-generating tools may discourage students' motivation to write and think on their own. It is worth reiterating that the learning activities in this subject seek to develop their cognitive skills and that these guidelines seek to prevent some AI tools from becoming automated plagiarism mechanisms.
- Thirdly, because of the risk that users, consciously or unconsciously, treat AI behavior as if it were human (Eliza Effect). For example, LLM-based tools do not "understand" the texts they produce; they mimic language patterns by synthesizing large volumes of data based on which they generate word sequences. This problem of anthropomorphizing machines is exacerbated by the fact that some chatbots have produced violent or harassing responses that could lead people to engage in harmful behavior.
- Fourth, because of the ethical and human rights implications associated with the use of certain AI systems given that: some tools tend to reproduce or amplify derogatory and discriminatory stereotypes associated with gender, race, ethnicity or disability; technologies could have been developed from massive copyright infringement; new forms of colonialism through the non-consensual extraction of information from historically marginalized communities; some tools would have been developed in contexts of labor exploitation; the development and operation of such systems generates a considerable carbon footprint; and, the potential violation of privacy and personal data protection rights of those who use them. For more information on the harms associated with the use of Generative AI, I recommend consulting the report “Ghost in the machine: Addressing the consumer harms of generative AI” (2023) by the Norwegian Consumer Council and the report

“Generating Harms: Generative AI’s Impact & Paths Forward” by the Electronic Privacy Information Center

Other resources for teachers

- Other guidelines on the use of AI at the university. I recommend consulting guidelines, policies, and principles of other universities, professors, and other organizations, such as: University of Washington, Hertie School, University of Tartu, Yale University, University of Helsinki, Technical University of Munich, University College London, UNESCO, UK Department of Education, Boston University, Warwick University, Universidad de Buenos Aires (IALAB), Universidad del Desarrollo, Russell Group, and Lance Eaton's list.
- Introductory literature on LLMs.
 - (A) For an introduction on how LLMs were developed and how they work (in particular, GPT-3), I recommend this text by Timothy B. Lee and Sean Trott.
 - (B) For an explanation on how LLMs work and what are their main risks, I recommend the text “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” by Emily M. Bender, Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell.
- Teaching resources.
 - (A) For those interested in the challenges that LLMs create for teaching writing in universities, I suggest consulting the paper entitled “Adapting College Writing for the Age of Large Language Models such as ChatGPT: Some Next Steps for Educators” by Anna Mills and Lauren M. E. Goodlad.
 - (B) I also suggest the materials contained in “AI Text Generators and Teaching Writing: Starting Points for Inquiry” by Anna Mills.
 - (C) The article by Anna Mills, Maha Bali and Lance Eaton entitled “How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process” documents and reflects on how professors from different universities around the world have adopted new pedagogical practices to respond to the challenges generated by Generative AI.
 - (D) MIT offers for free its DAILY Curriculum program that contains materials to teach basic AI skills to teenagers.
- Critical approach to the use of AI in academic settings.
 - (A) I recommend consulting Maha Bali's list of resources on developing critical literacy in AI.
 - (B) I also suggest reading a reflection by Mohammad Hosseini, Lex Bouter, and Kristi Holmes in which they argue that it is not desirable to adopt AI tools in education without first reflecting on their effects and biases and without having adopted measures to mitigate their risks.
 - (C) Naomi S. Baron's article “How ChatGPT robs students of motivation to write and think for themselves” discusses

how the use of LLMs can negatively affect students' creative processes. (D) Dimitrinka Atanasova's blog post reflects on how LLMs can reduce barriers for students and researchers to write in English (when this is not their first language), but that their effects on cultural inclusivity in higher education are more ambiguous. (E) Finally, I recommend the Kathryn Conrad's text "A Blueprint for an AI Bill of Rights for Education" that proposes rights for teachers and students that can inspire policy formulation and protective measures.

Guidelines open to change

- AI is a rapidly changing set of tools, which is why these guidelines will remain open to future evaluations, modifications, and revisions. • Throughout the academic semester, the professor will open spaces to discuss these guidelines with the students and, if necessary, modifications may be introduced through co-creation exercises. • I am grateful to my colleagues and others who provided comments and criticisms on initial versions of this document and also to those who have participated in the talks and workshops I have given on the use of large-scale language models in university contexts.

Changes introduced in this version of the guideline

- I introduced changes of form and substance with respect to the previous version of this instrument (v. 4.3), which I summarize below. • Changes of form: (A) The previous version was called "Guidelines for the Use of Artificial Intelligence in University Courses", the new title is broader since the instrument is useful for academic areas that go beyond classroom activities (e.g., research activities). (B) I introduced new sources that can be consulted in the form of hyperlinks throughout the text. • Substantive changes: (A) In point 1 of the new version of the guidelines I included a glossary of key terms and in point 3 I introduced a new sub-section with plausible and nonrecommended uses of LLMs by students. (B) In addition, in this version I included a new section of resources for teachers that includes examples from other guidelines, basic literature on LLMs, resources for teaching, and critical literature on the use of AI in educational contexts. (C) On the other hand, the previous version of this instrument indicated that the teacher would use synthetic text detection systems, but I have removed those allusions in the present version because the currently available systems produce a high level of false positives and negatives. For this

reason, for example, the Teaching Center at the University of Pittsburgh issued a press release stating that it did not endorse or promote any of the tools currently available

Use of these guidelines in your classes

- These guidelines are published under a Creative Commons Attribution 4.0 license, so if you are an educator interested in formulating your own policies or guidelines, you can share and adapt these guidelines as long as you make the corresponding attribution and indicate if you have made any changes.
- If you would like to send me your comments, criticisms, and suggestions, please write to me at juagutie@uniandes.edu.co

34. Editorial Policy, Publication Ethics and Malpractice Statement

Editorial Policy

The scientific and academic journal system of Pontificia Universidad Javeriana is composed of a set of indexed periodic publications managed jointly by the editorial teams of the Faculties and Editorial Javeriana. We are committed to open and transparent communication, scientific integrity and ethical standards in all aspects of the editorial and dissemination process. Our editorial policy is governed by the following principles:

Quality: It is our main goal to publish high-quality research that is relevant, visible and impactful. All submissions are subjected to rigorous peer-review by independent experts in the field. We strive to provide constructive feedback to authors to help them improve their work.

Ethics: We are committed to upholding the highest standards of ethical conduct in our editorial process. We require authors to adhere to ethical guidelines for research, informed consent, protecting privacy and confidentiality, and avoiding plagiarism and fabrication of data. Every paper is checked by means of anti-plagiarism and AI-generated text software detection.

Openness: We believe in open and transparent communication and support open access to scientific knowledge. We encourage authors to share their data and code whenever possible, and we require authors to disclose any potential conflicts of interest.

Diversity and inclusion: We recognize the importance of diversity and inclusion in research and editorial processes. We strive to ensure that our journals represent a broad range of origins, perspectives and experiences, and we are committed to fostering a culture of respect, inclusivity, and equity in our editorial policies and practices.

Innovation: Innovative and unconventional approaches to research and policy analysis are welcome. We encourage authors to challenge existing paradigms and propose new ideas for Latin America and the world. We offer traceability of the impacts of research in academia, industry and society. It is possible to include podcast, video, animations, or interactive content in the HTML versions of the articles.

Engagement: Seeking to engage with our readers, authors, reviewers, and stakeholders in an ongoing dialogue about the most pressing issues in Latin America and the world, we encourage authors to communicate their research findings in a clear and accessible way, providing a forum for discussion and debate around important issues.

Sustainability: As open access electronic journals, sustainability is promoted in all aspects of our operations. We recognize that our activities have an impact on the environment, and we strive to minimize this impact by adopting sustainable practices:

- Energy Consumption: minimizing our energy consumption by implementing energy-efficient practices in our office space, server rooms, and data centers is one of our main goals. We encourage our staff to switch off computers and other electronic devices when not in use. We use energy-efficient equipment wherever possible.
- Paperless Communication: We strive to operate as a paperless office by minimizing the use of paper and encouraging the use of digital communication, storage, and archiving methods. We also minimize printing and use recycled paper when printing is necessary.
- Green Hosting: We use web hosting services that are powered by renewable energy and strive to minimize the environmental impact of our online presence. We also optimize our

website and use selected directories, databases and scholarly indexes to reduce our carbon footprint.

Copyright and Licensing

The journals published by Pontificia Universidad Javeriana are registered under Creative Commons Attribution 4.0 International Public License. Thus, all articles may be reproduced, distributed, and publicly shared in digital format, as long as the names of the authors and Pontificia Universidad Javeriana are acknowledged. Others are allowed to quote, adapt, transform, auto-archive, republish, and create based on this material, for any purpose (even commercial ones), provided the authorship is duly acknowledged, a link to the original work is provided, and it is specified if changes have been made. Pontificia Universidad Javeriana does not hold the rights of published works and the authors are solely responsible for the contents of their works; they keep the moral, intellectual, privacy, and publicity rights.

Approving the intervention of the work (review, copy-editing, translation, layout) and the following outreach, is granted through an use license and not through copyright transfer. Publishing in this journal system does not generate direct royalties for contributors.

Privacy Statement

The names and email addresses submitted to the journals edited by Pontificia Universidad Javeriana will be used for the purposes stated solely and will not be made available for any other purpose or to any other party.

In accordance with Colombian Statutory Law 1581 of 2012 on Data Protection and related regulations, the Owner is informed that, by voluntarily providing and registering their data through the channels enabled on this website, they authorize their incorporation into a database under the responsibility of the Pontificia Universidad Javeriana, to be treated for the purpose of carrying out associative, cultural, recreational, sports, and social activities, management of social media and/or editorial content, historical, scientific or statistical purposes, administrative procedures and publications, and accounting, tax, administrative and advertising management and commercial

prospecting - Own advertising. Likewise, I authorize my personal images to be published in printed and audiovisual media, institutional websites and social networks.

It is optional to provide information about Sensitive Data, understood as those that affect privacy or generate any type of discrimination, or about minors.

The policy on the treatment of the Owner's data, as well as substantial changes that occur in it, can be consulted through the following email address: usodedatos@javeriana.edu.co. Likewise, it will be kept updated on the entity's website: [Agreement 657 \(in Spanish\)](#)

You may exercise your rights of access, correction, deletion, revocation, or claim for infringement of data, by sending a written request to Pontificia Universidad Javeriana, to the email address usodedatos@javeriana.edu.co, indicating in the subject the right you wish to exercise, or by ordinary mail sent to the address: Carrera 7 # 40-62 in the City of Bogota, Colombia.

No Author Fees

Our Diamond Open Access journal system is completely fee-free for authors. We believe that access to knowledge should not be limited by financial barriers, and we are committed to making our journals accessible to researchers and scholars from all backgrounds and regions.

We do not charge any submission fees, article processing fees, or publication fees. All articles published in our journals are freely available to readers worldwide, ensuring that research is disseminated widely and can be accessed by anyone with an internet connection.

Our commitment to no fees does not compromise the quality of our publication. We employ rigorous peer-review processes to ensure the integrity and validity of the research we publish. Our editors and reviewers are experts in their fields and uphold the highest standards of academic excellence.

We welcome submissions from all researchers, regardless of their funding or institutional affiliations. By eliminating financial barriers, we hope to encourage a more diverse and inclusive academic community and foster the growth of new ideas and perspectives.

Publication Ethics Statement

In order to maintain the highest ethical standards in all aspects of our operations, we have established the following ethics statement to guide the behavior of our editorial team, authors, and reviewers:

ETHICAL STANDARDS FOR EDITORS

Our editors are responsible for ensuring that all published content adheres to the highest ethical standards. As such, they will ensure that:

- The editorial decision-making process is objective, transparent, and free from any commercial or personal interests.
- Any conflicts of interest, whether actual or perceived, are identified and disclosed to the relevant parties.
- All submitted manuscripts are treated confidentially and any information related to the manuscript is kept confidential.
- Any suspected ethical misconduct or malpractice is investigated thoroughly, and appropriate actions are taken.
- The editorial team takes measures to ensure that any published research involving human subjects, animals, or potentially hazardous materials adheres to relevant ethical guidelines.

ETHICAL STANDARDS FOR AUTHORS

Authors are responsible for ensuring that their research adheres to the highest ethical standards. As such, they will ensure that:

- The research presented in the manuscript is original and has not been published elsewhere.
- The manuscript is free from plagiarism, falsification, or fabrication of data.
- Any potential conflicts of interest are disclosed in the manuscript.
- Any research involving human subjects, animals, or potentially hazardous materials adheres to relevant ethical guidelines, and the necessary permissions have been obtained.

- The manuscript is written in clear and concise language, and the research is presented in an accurate and unbiased manner.

ETHICAL STANDARDS FOR REVIEWERS

Reviewers play a critical role in maintaining the quality and integrity of our journals. As such, they will ensure that:

- The review process is objective, transparent, and free from any commercial or personal interests.
- Any conflicts of interest, whether actual or perceived, are identified and disclosed to the editorial team.
- Any suspected ethical misconduct or malpractice is reported to the editorial team.
- The review is conducted in a timely and constructive manner, and any criticisms or suggestions are presented in a respectful and professional manner.

COPE adherence Statement

Our scholarly journals are committed to upholding the highest ethical standards in publishing and adhere to the guidelines and best practices set forth by the Committee on Publication Ethics (COPE):

- Editorial Independence: Our editorial decisions are independent and free from undue influence, whether from owners, sponsors, advertisers, or other third-party entities.
- Peer Review: Our peer review process is transparent and impartial, and all manuscripts are subject to rigorous review by qualified experts in the field.
- Authorship and Contributorship: We follow the Contributorship Defined (CreDIT) model for authorship, which acknowledges the contributions of all individuals involved in the research process.
- Conflict of Interest: We require authors, reviewers, and editors to disclose any potential conflicts of interest that may compromise the integrity of the research or publication.
- Research Misconduct: We do not tolerate research misconduct, including plagiarism, fabrication, and falsification of data.

- Data and Reproducibility: We require authors to provide all relevant data and materials necessary to reproduce their findings and support their conclusions.
- Intellectual Property: We respect the intellectual property rights of authors and require authors to obtain appropriate permissions and/or licenses for any copyrighted material used in their manuscript.
- Post-publication Discussions and Corrections: We provide a platform for post-publication discussions and corrections, and we are committed to correcting any errors or inaccuracies in our publications.

By adhering to the guidelines and best practices set forth by COPE, we are committed to promoting transparency, integrity, and responsible publishing practices in all aspects of our journals. We encourage our authors, reviewers, and readers to report any concerns or potential breaches of these practices to our editorial team for review and action.

6.1. AUTHORSHIP AND CONTRIBUTORSHIP

Our scholarly journals follow the Contributorship Defined (CreDIT) model for authorship, which acknowledges the contributions of all individuals involved in the research process. The CreDIT model promotes transparency and accountability in authorship and helps ensure that all contributors are recognized for their contributions.

In accordance with the CreDIT model, authors must explicitly state their individual contributions to the research presented in the manuscript. The contributions may include, but are not limited to, the following:

- Conceptualization: This includes the formulation of research questions, study design, and hypothesis development.
- Methodology: This includes the development or adaptation of methodologies and protocols for data collection and analysis.
- Data Curation: This includes the management and organization of data, including data entry, quality control, and data archiving.
- Formal Analysis: This includes the application of statistical methods and techniques to analyze the data.

- Investigation: This includes the collection and acquisition of data and/or samples.
- Resources: This includes the provision of funding, facilities, equipment, and other resources necessary for the research.
- Writing - Original Draft: This includes the initial drafting of the manuscript, including the writing of the introduction, methods, and results sections.
- Writing - Review & Editing: This includes the critical review and editing of the manuscript, including the revision of the introduction, methods, and results sections.
- Visualization: This includes the creation of visual aids, figures, and tables to present the data.
- Supervision: This includes the oversight and guidance of the research team and/or project.
- Project Administration: This includes the coordination and management of the research project, including obtaining ethical approvals, obtaining funding, and other administrative tasks.

Each author must provide a statement outlining their individual contributions to the research presented in the manuscript using the above categories. All authors must agree on the final author list and order of authors.

Our journal system recognizes that authorship credit should be based on substantial contributions to the research and does not allow honorary or gift authorship. We also encourage authors to acknowledge the contributions of individuals who do not meet the criteria for authorship but have contributed to the research in other ways.

By following the CreDIT model for authorship, we aim to promote transparency and accountability in authorship and recognize the contributions of all individuals involved in the research process.

6.2. ARTIFICIAL INTELLIGENCE GENERATED TEXT AND IMAGES

Artificial Intelligence (AI) generated text and images are not prohibited, but must be used in a responsible and ethical manner. All users of these technologies, like ChatGPT or Midjourney must adhere to the following principles:

- Respect for Human Dignity and Rights: All text generated text and images must respect the dignity and rights of all individuals, regardless of their race, gender, sexual orientation, religion, or any other characteristic. Hate speech, discrimination, or any other form of disrespect towards individuals or groups is strictly prohibited.
- Accuracy and Fact-Checking: Users must take responsibility for verifying the accuracy of any text generated before relying on it. Users must take steps to ensure that the generated text is accurate and reliable.
- Avoiding Harm: Users must ensure that any text generated does not cause harm to individuals or society as a whole. This includes avoiding the dissemination of false or misleading information that could lead to harm, as well as avoiding the promotion of harmful behaviors or activities.
- Transparency: Users must be transparent about the use of Artificial Intelligence and any text or images generated by the model. This includes acknowledging the use of the software when publishing or sharing AI generated text or images, and providing clear explanations of how the content was generated.

AI software attributions must not appear as author or co-authors of any submission, nor the bibliographic reference list, but in the Methodology and/or Acknowledgements sections.

6.3. ORIGINALITY AND PRIOR PUBLICATION

The manuscripts submitted to this Journal System should be original. Authors must appropriately cite and attribute any work that has influenced their current contribution. This includes both published research articles and unpublished work such as conference presentations, theses, dissertations, multimedia or alternative sources.

Submissions should have not been previously published or submitted for publication elsewhere, in whole or in part in any language. Authors should declare that the manuscript is not a duplicate of any previously published work.

We do not consider posting on a preprint server and theses or graduation papers on institutional repositories (green-access portals) to be duplicate publication, because these typologies are not evaluated by peer-review process and have not received editorial or publishing intervention.

- If there is a preprint version, authors must cite it in the new manuscript submitted to the journal (including Digital Object Identifier DOI and/or URL address) to ensure there will be a link to the original record.
- After publication, the following text can be added to preprint versions to encourage readers to use and cite the final published version of the article:

"This is an original manuscript of an article published in [JOURNAL TITLE], available online: <https://revistas.javeriana.edu.co/> | [Article DOI].".

All submitted manuscripts will be checked for originality using plagiarism detection software. Any text similarity detected above 30 % may lead to further evaluation to determine if there is text recycling or any other malpractice.

6.4. CONFLICTS OF INTEREST

To uphold the highest standards of integrity and ethical conduct in all aspects of our publishing activities, all editors, reviewers, and authors are required to disclose any potential conflicts of interest that may arise during the publication process.

Conflicts of interest can arise in a number of different forms. For example, an editor may have a financial or personal relationship with an author or organization whose work they are evaluating. Similarly, a reviewer may have a personal or professional connection to an author that could influence their assessment of the manuscript.

To ensure that conflicts of interest are identified and appropriately managed, we require all editors, reviewers, and authors to disclose any relevant financial or non-financial relationships that could be perceived as having the potential to influence their judgment. This includes any relationships with organizations, companies, or individuals that may have a direct or indirect interest in the research or findings presented in the manuscript.

In cases where a potential conflict of interest is identified, we will take appropriate steps to manage the conflict and ensure that the integrity of the publication process is maintained. This

may include assigning a different editor or reviewer to the manuscript, or requiring additional disclosures from the authors.

6.5. DATA SHARING AND REPRODUCIBILITY

Data Sharing Policy: authors are encouraged to make all data underlying their research available to other researchers, either through open data repositories or by providing the data upon request.

Data Accessibility: Authors should provide information on how to access their data, including any necessary software or programming code required to reproduce their results. The data should be accessible and usable by other researchers, and should be accompanied by clear documentation that describes the data and its provenance./p>

Data Citation: Authors should cite all datasets used in their research, and should provide complete references to these datasets in the manuscript. Data should be considered a valuable scholarly output and should be cited and recognized as such.

Reproducibility: Authors should provide all necessary information to enable other researchers to reproduce their research, including detailed descriptions of their methods, software code, and data sources. Authors should also include detailed descriptions of any data preprocessing, analysis, or interpretation performed in their research.

Quality Control: Authors should provide information on any quality control procedures used in their research, including how they identified and dealt with outliers, missing data, or other issues that may affect the validity of their results.

Open Science: Open science practices are recommended, such as pre-registration of research protocols and the use of open-source software tools. Open science practices can help to increase transparency, reproducibility, and accountability in research.

6.6. MALPRACTICE STATEMENT

To maintain the integrity of the research process and ensure high quality in every published work, scholarly journals must address various malpractices that can occur during the publication

process. Some of the most common malpractices that scholarly journals encounter are plagiarism, salami slicing, double submission, and data fabrication/falsification:

- *Plagiarism* occurs when an author uses someone else's work, ideas, or words without proper attribution or citation. This includes copying and pasting text, paraphrasing without proper citation, or presenting someone else's work as one's own. Plagiarism undermines the credibility of the research and violates ethical and professional standards.
- *Salami Slicing* is the practice of breaking up one study into multiple smaller studies to increase the number of publications. This practice can lead to the duplication of data and ideas, and can also result in incomplete or misleading information being published.
- *Double Submission* occurs when an author submits the same manuscript to more than one journal at the same time. This practice can waste the time and resources of multiple journals, as well as delay the publication of other research.
- *Data Fabrication/Falsification* occurs when an author knowingly or unknowingly fabricates or falsifies research data or results. This can include manipulating data to obtain desired results, or fabricating data that was never collected or analyzed. Data fabrication/falsification undermines the credibility of the research and can have serious consequences for the scientific community and the public.

Other malpractices that can occur in scholarly publishing include conflicts of interest, duplicate publication, citation manipulation, and guest authorship. All of these malpractices can have serious consequences for the credibility of the research, the reputation of the authors, and the integrity of the scientific community.

6.7. ALLEGATIONS OF RESEARCH MISCONDUCT

In the event that ethical misconduct or malpractice is suspected, the editorial team will conduct a thorough investigation. If the investigation confirms the misconduct or malpractice, appropriate actions will be taken, including, but not limited to:

- Contacting the relevant authorities, such as the author's institution or funding agency.
- Retracting the published article.
- Banning the author from submitting to our journal system.

- Removing the author from the editorial board.

We take any suspected ethical misconduct or malpractice very seriously, and we will take all necessary steps to ensure that the integrity of our journals is maintained.

6.8. CORRECTION AND RETRACTION GUIDELINES

Here are some general guidelines for authors who need to correct or retract a scholarly journal article:

- Act quickly and responsibly: If you discover errors or misconduct in your article, it is important to act quickly and responsibly. Delaying correction or retraction can cause harm to the scientific community and damage your own reputation.
- Contact the journal: You should contact the journal's editor or publisher as soon as possible to inform them of the error or misconduct. Explain the situation clearly and provide any evidence or documentation that supports your decision to correct or retract the article.
- Provide a clear reason for the correction or retraction: You should provide a clear and detailed explanation of why you are correcting or retracting the article. This can include minor errors, errors in the data, methodology, or analysis, as well as ethical or legal concerns.
- Include a correction or retraction notice: The journal will likely require you to submit a retraction notice to be published alongside the original article. The retraction notice should include the article title, authors, and publication details, as well as a clear and concise explanation of why the article is being corrected or retracted.
- Correct the record: If the retraction is due to errors in the data or analysis, you should work with the journal to correct the record by publishing a correction or erratum.
- Be transparent and accountable: Retracting an article can be a difficult and embarrassing process, but it is important to be transparent and accountable. You should take responsibility for the errors or misconduct and work to prevent similar issues in the future.
- Follow ethical guidelines: Retracting an article can have serious consequences for the scientific community, so it is important to follow ethical guidelines and best practices. You should be honest, transparent, and respectful throughout the retraction process.

Retraction

Retraction should be a last resort and should only be used in cases of serious errors or misconduct. If you are unsure about whether to retract an article, you should seek advice from the journal's editor, a colleague, or a professional association. By taking appropriate steps to retract flawed or unethical research, journals can help maintain the integrity of the scientific record and protect the trust of their readers and the broader scientific community.

Retractions can also be issued by the journals or Editorial Pontificia Universidad Javeriana. The following guidelines outline the steps that journals should take when retracting a published paper:

- Identify the reason for the retraction: Journals should clearly state the reason for the retraction in the retraction notice. Common reasons for retraction include scientific errors, plagiarism, data manipulation, and ethical violations.
- Notify the authors: The journal should contact the authors of the retracted paper to inform them of the decision to retract and the reason for the retraction. The authors should be given an opportunity to respond and provide any additional information that may be relevant.
- Publish a retraction notice: The journal should publish a retraction notice that clearly identifies the retracted paper, the reason for the retraction, and any relevant details. The notice should be prominently displayed in the journal and clearly linked to the original article.
- Provide a detailed explanation: The retraction notice should provide a detailed explanation of the reason for the retraction, including any specific errors or ethical concerns. The notice should also include a statement about the impact of the retraction on the scientific record and any potential implications for future research.
- Correct the literature: The retraction notice should include a statement about any corrections or clarifications that may be needed in other publications that cited the retracted paper. The journal should also work with other databases and indexes to ensure that the retraction is properly reflected in the literature.
- Follow established guidelines: The journals edited by Pontificia Universidad Javeriana follow established guidelines for retraction developed by the Committee on Publication Ethics (COPE). These guidelines provide a framework for handling retractions and ensuring transparency and consistency in the process.

- Consider sanctions: In cases where there is evidence of misconduct, the journal may need to consider imposing sanctions on the authors, such as banning them from submitting future papers to the journal or reporting the misconduct to relevant institutions.

Double-Blind Peer-Review Process

Our scholarly journals follow double-blind peer review policy to ensure fairness and impartiality of the review process. This policy ensures that both the reviewers and authors remain anonymous to each other throughout, thus reducing the potential for bias and promoting objectivity:

Reviewer Anonymity: Reviewers are not provided with the name, affiliation, or any identifying information about the authors of the article being reviewed. This ensures that the review is based solely on the quality and merit of the research presented in the article.

Author Anonymity: Authors are not provided with the name or affiliation of the reviewers. This ensures that the authors' opinions of the reviewers do not influence the submission and review process.

Reviewer Selection: Reviewers are selected based on their expertise and experience in the field of the article being reviewed. Our journal's editorial board selects reviewers who have no conflict of interest with the authors of the article.

Review Criteria: Reviewers are asked to evaluate the article based on the following criteria:

- Originality of the research
- Significance and relevance of the research to the field
- Clarity and coherence of the argument and presentation
- Methodological rigor and accuracy of the data presented
- Appropriate use of references and citations

Reviewer Feedback: Reviewers are asked to provide constructive feedback to the authors on how to improve the quality and clarity of their research. Reviewers are encouraged to be critical but constructive in their feedback and to provide specific suggestions for improvement.

Editorial Decision: The editor-in-chief makes the final decision on whether to accept or reject the article for publication. The editor takes into account the reviewers' comments and feedback, as well as the article's originality, significance and quality.

Reviewer Confidentiality: All reviewers are required to maintain the confidentiality of the review process and not to share the content of the article or their review with anyone else without the editor's permission.

Open Access and Self Archiving Policy

This scientific and scholarly journals system works under a Diamond Open Access model. This means, it does not require any publication, processing, access or subscription fees. Instead, it is entirely funded and supported by Pontificia Universidad Javeriana.

Postprint and self archiving are allowed (to post the published version anywhere) with no embargo under the following conditions:

- Published source must be acknowledged
- Must link to the original work
- Must indicate if changes have been made

The Publisher SHERPA/RoMEO license can be accessed in the following link:
<https://beta.sherpa.ac.uk/publisher/1759>

Editorial Pontificia Universidad Javeriana is part of the Public Knowledge Project Preservation Network (PKP PN) to deposit and preserve all contents published by using the LOCKSS program.

Advertising and Marketing

Editorial Pontificia Universidad Javeriana does not regularly invest in advertising for the Scientifia and Scholarly Journal System. However, it ensures that any possible advertising piece does not interfere with the integrity and quality of the published contributions. Our advertising and marketing policy statement includes the following guidelines:

- Advertisements do not compromise the quality or integrity of our publication or compromise our independence as a scholarly journal.
- Advertisements are not misleading or deceptive, and should comply with all applicable laws and regulations.
- Advertisements do not imply endorsement or sponsorship by our journals or our editorial teams.
- Advertisements are clearly identified as such and are visually distinguishable from editorial content.
- Advertisements are targeted to relevant and appropriate audiences and not interfere with the user experience of our readers.

We reserve the right to reject or remove any advertisement deemed inappropriate, misleading, or inconsistent with our values and mission.

We believe that responsible advertising and marketing practices can help promote the visibility and impact of our journals and increase the visibility and accessibility of the research we publish.

35. Readiness of the judicial sector for artificial intelligence in Latin America

INDEX PREFACE AI, Justice and Digital Transformation Policies in the Latin American Public Domain CASE ARGENTINA Assessment of the readiness of the judicial system for the adoption of artificial intelligence CASE CHILE Artificial Intelligence and The Judiciary. Chile and its Pending Challenges. CASE COLOMBIA Analytical and exploratory framework CASE MEXICO Readiness assessment for the adoption of Artificial Intelligence in judicial systems CASE URUGUAY Preparedness assessment for the adoption of Artificial Intelligence in judicial systems of the region

Justice and Artificial Intelligence:

Conceptual foundations of the research The development of Artificial Intelligence (AI), especially in its forms related to automatic learning or machine learning, has become one of the most relevant challenges faced by contemporary societies. Its impact on different sectors is being assessed in a context of skepticism, dystopia and optimistic narratives, which may even misrepresent the

possibilities and limitations of this technology. While the most advanced AI has become a pervasive, widely discussed element in the media in fields such as recruiting and marketing, it is permanently going forward in areas such as healthcare, transport and education (Pombo et al, 2020). This AI progress in different types of activities has promoted in the last years an emerging, albeit growing, discussion in the judicial domain as well. This sector has some characteristics that make it particularly appealing for the implementation of AI. Most notably, the great volume of information and data generated in the administration of justice render it a particularly relevant space to implement AI techniques that make it possible to systematize, deduce, and create patterns and prediction in less time and with greater resource efficiency. The introduction and implementation of AI and its advanced expressions based on machine learning, and in particular its sub-field known as deep learning, pose varied challenges for the legal systems in any jurisdiction (Chen et al 2019). There are distinctive characteristics of the operation of the Judiciary and its regulatory and cultural environments that should be understood in the context of an increasingly more widespread discourse regarding the inclusion of machine learning techniques in the different Latin American judicial systems. Firstly, and at the most basic level, technology is helping inform, support and advise the people involved in the justice system in the form of supporting technology. Secondly, AI systems may replace functions and activities that used to be performed by humans, which would be called “substitution technologies”. Lastly, at a third level, technology may change the way judges work and provide very different forms of justice, acting here as a disruptive technology, especially when processes change significantly and the predictive analysis may reshape the allocation function. At the second and third levels is where the main questioning related to the impact of technology on judges’ roles and functions regarding the allocation function is raised (Sourdin, 2018). This diagnosis is also shared with another report drafted at the European Commission that we are going to deal with later on. The introduction of AI in the judicial domain may have a transformation impact, with risks and benefits for a better governance of the legal system. As for the positive aspects, the use of AI in this field may have an influence on the internal organization of these systems, as well as on the administration of justice for citizens of these countries, making it more affordable, accessible, transparent and agile. It may improve accountability and efficiency, and reduce the workload (COE, 2019). At the same time, there are risks associated with the misuse of these systems based on the lack of knowledge of its operation as well as the lack of responsibility of decision-makers, designers of the technology

architecture, as there are problems inherent to the design and operation of these systems that revolve around accountability, equity, access, transparency and interpretability, many of which have been considered in practice and in the recent literature as ethical problems. These problems go beyond the use of AI systems in judicial realms and are more general for their implementation in different sectors. However, the criticality in a legal structure in democratic systems and for the full exercise of human rights deserves an even more careful consideration of the inclusion of AI systems in justice. As there is not a unique approach on AI as a tool to provide solutions for the challenges faced by the Judiciary in different countries of the region, nor is there evidence that its use in this sector is questioned, while there is no universal judicial model, this work aims at understanding the direction of discussions and practices, policies and regulations, many of them incipient, about the implementation of AI systems in the Judiciaries in the region. This work seeks to complement international discussions as well as bridge existing gaps about the region by means of empirical evidence on the state of the issue in AI systems in legal structures, considering the political, institutional and digital technologies adoption design. To that end, this work develops an analytical framework that makes it possible to map the national approach to this topic, taking into account the national ecosystem and the international macro level; the meso level, adjusted to the dynamics inherent to the judicial system as organizations; and the micro level, related to concrete players and institutional entrepreneurs in this subject matter. Given the current context of uncertainty, skepticism, magnified expectations and economic interests of providers and a flourishing industry, this study points at filling the knowledge void related to the impact of AI and its meaning for the judicial experience in the Latin American context. This framework takes six dimensions that consider the reference indices, regulations and the strategy related to AI; the governance of the judicial updating processes; the diagnosis and the capacities to adopt AI; the existing conditions for the deployment of AI in the legal domain; and the sense of opportunity that AI entails for the judicial domain in each country. These six dimensions are addressed with 50 indicators, aspects which will be developed in the second part of this report.

History of AI updating, digitalization and inclusion, and links with the judicial sector

Digital technologies based on the internet are increasingly perceived as an instrument to make justice go forward. When properly implemented, they can enable access, reduce differences and

serve a wider audience. Technology makes it possible to provide services overtaking traditional methods and improving the experience of the population when they resort to the Judiciary. Accessibility, transparency and efficiency are principles that may be derived from a conscientious and responsible implementation of technology to serve justice (COE, 2019). A balanced implementation of technology may help achieve a people-centered justice. In this sense, the Sustainable Development Goals (SDG), in its item 16.6, have explicitly addressed the challenge of “creating efficient and transparent institutions at all accountable levels”. Accountability is still a key priority to administer and impart justice, and AI implementations promise to achieve it by getting rid of the inherent human bias. However, many algorithmic techniques used for deep learning are based on “black box” approaches, which do not allow for accountability. In civil courts, it is unimaginable to meet the objective of the SDG 16.6 of failing to solve the delays affecting all the cases (even the simplest ones in most contexts) and to move forward towards effective solutions –following the Rule of Law Index (World Justice Project 2019). Based on the analysis of the data coming from the countries that are part of this study, postponements and delays are endemic in the bureaucratic organization in justice systems. This problem has grown and become more evident in the context of the COVID-19 pandemic. As a general purpose technology with diverse techniques and implementations that go from the automation of processes to tools emulating human intelligence, the scope of AI is ample, and the needs, as well as the institutional capacity and preparedness to include these technology sets, also depend on the existence and availability of other technical conditions: network connectivity, computers and information systems, digitalized documents and the like, to name a few. Now then, by and large, AI and Information and Communication Technologies (ICTs) themselves do not explain why and how change takes place in public service organizations (OECD, 2019). Thus, the concept of digital transformation emerges as a necessary condition to assess the capacity and preferences of judicial institutions to incorporate digitalization and its processes. Even though the speed and broadness of changes in public services through the adoption of digital technologies, including the administration of justice, have been in the agenda for almost two decades (Heeks, 2001), the current stage is evaluated from the concept of digital transformation and “disruptive technologies”, rather than “electronic government”, as was coined at the beginning of this century. Digital transformation entails institutional options that are neither lineal nor deterministic by the available technology (Filgueiras et al, 2019). Technologies and institutional processes interact to put

together a complex framework of digital transformation (Fountain, 2001). In particular, the digital transformation based on AI, as well as the implementation of AI in the entire public sector, presents a special condition that shows the double relationship of the use of this technology for governments: on the one hand, the obligation of States to protect citizens against possible algorithmic damages strains with their temptation to increase their own efficiency. In other words, the balance between governing algorithms at the same time as it governs by algorithms (Kuziemski, Misuraca, 2020: 2). The introduction of AI is framed by several discussions about changes in the judicial environments, starting with the nature of the adjudicative function of judges, as well as the context in which lawyers, courthouses and other sectors are currently using technology (Sourdin, 2018). From the perspective of public policies of digital transformation and in line with Sourdin (2018), the environment, the types of provisions, terms, rates, the provision and processes of user identification affect the institutional choices that opt for an interface or digital solution. Thus, the digital transformations of public services depend on institutional processes with a variety of preferences for digitalization. These preferences point to choices of digitalization processes of certain services, but not others (Filgueiras, Cireno, Palotti, 2019). The adoption of digital transformation policies in any organization of the public sector allows for new forms of mediation and institutionalization of practices that represent a window of opportunities to promote inclusion, efficiency and effectiveness for the provision of services. Nonetheless, this same literature points out that a digital transformation process takes place within institutional contexts where choices and decisions are made based on players' different preferences. The challenge becomes evident if we add that these preferences are made up by their knowledge and understanding of the topic at issue. With such a complex and ubiquitous technology as AI, the need to provide greater understanding once it is incorporated in the information systems of the national judicial system becomes vital. The convergence of ethics as an AI governance instrument and approach has been a sustained element in the last five years, with the development of over 60 documents about ethical principles and considerations for the evolution of AI (Fjeld et al, 2019). However, the only specific instrument for the Judiciary was promoted by the European Commission for the Efficiency of Justice (CEPEJ) of the Council of Europe, and it was adopted by the end of 2018. It is the first European document that establishes the basic principles for the use of Artificial Intelligence in the national judicial systems, and it portrays both the fundamental values and the basic methodological requirements for the creation and development of algorithms.

Basically, these principles expressed in Europe have become some of the few guidelines that reflect on the connection between the AI and the Judiciary. These principles are: respect for human rights, that is to say, the design and implementation of AI should be compatible with the fundamental rights; the principle of non-discrimination, which prevents the development or intensification of discrimination processes towards individuals or groups; the principle of quality and safety, which refer to issues of data governance focusing on their quality and safety; the principle of “transparency, impartiality and equity”, which allows data processing to be accessible, auditable and understandable; and lastly, the users’ control, which makes it possible to avoid a prescriptive approach, and to ensure that users are informed players and have control over their decisions (COE, 2019). An additional aspect that the CEPEJ document highlights are the recommended uses of AI in judicial systems –with an European focus given its area of competence-, although they constitute a relevant basis for the development of future recommendations for the practices of the use of AI in Latin American judicial systems, as well as other regions. The report identifies four types of uses: those to be promoted, those that require considerable methodological precautions; those that should be developed following clear scientific guidelines; and, finally, those uses that should be considered with the utmost reservation. Each of them will be analyzed briefly below As part of the AI uses to be promoted are an improvement of the jurisprudence through natural language processing techniques and data visualization that enable the task of studying and looking for sources beyond searching with key words; additionally, techniques that make it possible to gain access to justice, such as chatbots, and the creation of new strategic tools related to the activity at courthouses, allowing for projections on budgets, personnel and key performance indicators. Those uses that call for considerable methodological precautions include applications that require help for the elaboration of scales and parameters in some civil litigation; support for alternative mechanisms to deal with disputes in civil cases; online mechanisms to deal with disputes; and the use of algorithms in criminal investigations to help in the identification of locations where criminal offenses take place. The third group of uses that should be considered is only additional scientific studies that include the profiling of judges and the anticipation in court decisions. Finally, according to the COE document, among the uses to take into account with the utmost caution are those that use algorithms in criminal cases to make profiles of individuals and regulatory quantifications. In the second case, these uses encompass the systematization and standardization of information regarding the set of decisions produced by

colleagues, with the risk that a judge's decisions may be predetermined or circumscribed by the set of data framed as precedents. This classification does not constitute a definitive guideline but a relevant mapping to establish the risk level in discussions and experimentations currently carried out in the region.

An exploratory analytical framework designed for the Latin American reality

There is a strong shortage of analytical instruments devoted to analyze the Artificial Intelligence capacities in state agencies, and a complete lack of instruments designed to enquire on institutional preferences about the subject matter. With regard to state capacities, the only analytical framework intended to assess the maturity regarding AI is the Government AI Readiness Index, developed by Oxford Insights and the International Development Research Centre (IDRC) in 2017 (it also has 2019 and 2020 editions). Now then, it is only an instrument intended for the analysis at a government level and, strictly speaking, its main basis is a kind of summary of other indices (such as the Networked Readiness Index or the UN e-Government Survey). Moreover, it is oriented towards the scoring of countries on a number scale. Listing countries in a ranking may be useful for an overall look about the state of the issue at a global level; however, when the intention is to delve deeper in the cases with the premise that through their details and contrasts we may learn more about them, their preferences and the problem under study itself, international indices do not offer the best heuristic alternative. Likewise, they tend to introduce in an underlying manner a lineal idea of the capacity maturity processes, in general, from the identification of a base line built around the development paths deployed by the main economies of the world (which have the resources to fund global research projects), biases that are later adopted as expected paths by the other countries (and which may be more or less controlled by researchers in each case). That is why we believe that is vital to enquire, in turn, about state preferences (for example, in terms of the institutional design of public policies or the adoption of different international cooperation standards), as well as the dynamics that explain their formation processes at a national level. This need is exacerbated when we enquire about phenomena such as AI use, where there is no universal standard but an emerging "regime complex". Regarding studies on state preference for AI use, however, the shortage pointed out for the capacity study only gets stronger. In its 2020 edition, the Government AI Readiness Index incorporated a sub-index which covers 34 countries, where

the concept of “responsible use” of AI is explored. What countries understand by “responsible use” is already an initial problem of the preference processes. However, we deem it necessary to go much further to elaborate a framework where both aspects, preferences and state capacities, are addressed in a comprehensive way. That said, when the subject under study is not the preparedness to incorporate AI in governments but in the Judiciary, such shortage becomes a complete absence of available analytical frameworks. That is why the project started off the elaboration of an original analytical framework that would make it possible to explore Latin American cases, which was achieved from the joint work of all the researchers involved in the project during three intense discussion sessions held remotely between August and October 2020, given the pandemic scenario. The main challenge that came up during that first stage was how to design an analytical framework that makes it possible to address really diverse cases given their differences in institutional (for example, there are federal and unitary countries, which has consequences on the institutional design of the judicial system) and socio-historical terms (there are countries with varying degrees of economic development, and different patterns of international integration), and in terms of digital transformation trajectory. In order to overcome this obstacle, we have created a consistent framework with six dimensions and over 50 empirical indicators, ranging from the general to the individual approach; moreover, they make it possible to explore both the precedents that are to frame the AI rollout as well as the existing developments. The dimensions considered were the following:

- General mapping of the case according to reference indices. This dimension intends to place the case in the universe, especially in terms of its trajectory regarding information development, digitalization of public services and openness for public data management. To that end, eight well-known international indices are considered as proxy, and they address different aspects regarding the problem at issue: the aforementioned Government AI Readiness Index and the e-Government Index of the UN, the Global Open Data Index designed by the Open Knowledge Foundation, the Network Readiness Index designed by the World Economic Forum, the index of the World Intellectual Property Organization, and the ICT Development Index drafted by the International Telecommunication Union.
- Even though this dimension is the first one addressed for the elaboration of each national investigation, in the reports we will find it as Annex I. This is so because none of the studies based on international indices addresses specifically the judicial sector, so reading them entails a complement for each national report. Now then, reading is recommended as it is useful as a fine introduction for the subject matter being analyzed.

Regulations and strategies about AI. This dimension is divided into three sub-dimensions: i) National regulation; ii) National strategies; and iii) International alignments in terms of regulatory standards. The regulations consider both specific rules about AI and those with a direct impact on AI governance, as well as rules about the Judiciary updating processes and the protection of personal data, and it explores the constitutional hierarchy of Human Rights agreements. The strategies consider both digital strategies as well as those specific for AI, whether they are related to the judicial domain or not. Finally, in terms of geopolitical alignments, the degree of adequacy to the European Data Protection Regulation, the endorsement of Convention 108 and the principles about AI use of the OECD are considered, as well as provisions related to AI in digital commerce agreements.

- Governance of judicial updating processes, with emphasis on AI. This dimension is divided into three sub-dimensions: i) Governance; ii) Players; and iii) Political discussion. In terms of governance, on the one hand, the Judiciary institutional organization outline is explored and, on the other, its outline of data and statistical management. Regarding players, firstly, the focus is on public agencies that take part in the judicial updating processes at national, sub-national and federal levels (the last one, where it is applicable), agencies and networks that have dabbled in the development of AI in the judicial domain, coordination mechanisms of issues related to the digital agenda in the judicial domain (with emphasis on AI use) and formal mechanisms of strategic coordination about AI use at state level. Secondly, the focus is laid on alliances and networks among the players mentioned. Finally, as for the political discussion, the forums and environments of the state sector in general and the judicial one in particular are considered, where discussions about AI use in the State are taking place.
- Diagnosis and development of capacities for the adoption of AI. This dimension is divided into three sub-dimensions: i) Digital judicial services; ii) Human talent; and iii) Institutional resources. In terms of Digital judicial services, the focus is on the developments regarding process digitalization, inter-operability and expert systems. Regarding Human Talent, the existing capacities of the judicial personnel regarding the use of complex digital tools, such as AI, are surveyed, as well as the training plans in the judicial career with official support that incorporate the topic. As for Institutional resources, the existence of a top state agency with AI leadership and appropriate budgetary resources, and a leading agency with an integral plan for the use of AI in the Judiciary is considered.
- Existing conditions for the deployment of AI in the judicial domain. This dimension is divided into six sub-dimensions: i) Starting point; ii) Theoretical basis; iii) Axiological basis; iv) Success metrics; v) Risk matrix; and

vi) Responsibilities. In terms of starting points, we refer to surveying the initiatives effectively existing to incorporate AI in the Judiciary (at any level), as well as its objective and need, that is to say, their justification and the objectives associated with them. Theoretical basis refers to the analysis of the concepts of “humanity” and “justice” used as a theoretical basis for concrete propositions around AI in the Judiciary. Meanwhile, the axiological basis examines the principles and values that AI is framed into, with emphasis on the considerations about ethics. The success metrics refer to metrics used in AI-based projects to assess results. The risk matrix refers to the risks considered by public policy makers as related to the use of AI in the Judiciary, as well as the criteria based on which the risks structures are built, and the safeguards designed to deal with the unwanted effects of the use of AI in the Judiciary. Finally, responsibilities refer to the provisions related to the allocation of responsibilities for AI use in the judicial domain. • AI as an opportunity for the judicial sector. This last dimension is divided into three sub-dimensions: i) Sense of urgency; ii) Applicability conditions; and iii) Reflection on the role of Justice. The sense of urgency is focused on the key challenges of the judicial system in the county that may be dealt with better through AI, and the reasons why the initiatives to incorporate AI may be an opportunity for the Judiciary, beyond the ones effectively existing. The aim of applicability conditions is to distinguish the areas of the Judiciary where it is and it is not advisable to move forward with the use of AI, as well as the issue of how appropriate the available AI techniques are for the problems to be solved in the Judiciary. Finally, the reflection on the role of Justice suggests exactly that: an analysis focused on the systemic and long-term impact that AI use may have on Justice for the role of judges.

Description of the five national reports The second stage of the research project, which took place between November 2020 and May 2021, was devoted to preparing the national reports. In April, two remote workshops were held to pool the preliminary versions. Each report received comments from at least two researchers from other research teams. The final versions of each report were prepared on such basis. As has been said, the same analytical framework was adopted as the basis for the five national reports. Now then, the general criterion established was for each research team to be able to adapt the framework wherever necessary, keeping the fundamental elements and justifying such methodological strategy, by virtue of the possibilities offered by each case (in terms of evidence of initiatives effectively existing and access to verification means). Likewise, the study admitted some leeway for styles diversity. As a result, each report is, at the same time, a unique

case study and an input to identify interest contrasts and similarities, without being a comparative analysis. Below we will point out the most interesting findings of each report. Later, the main similarities and differences among the five will be mentioned.

The study about Argentina, created by Gonzalo Bustos Frati and Bruno Gorgone (CETyS), is centered on what they call “effective initiatives of institutional innovation associated with AI in the judicial domain in Argentina”. What they mean by “effective initiatives” are the lines of action of a courthouse or another judicial player of the public sector that have been translated into the use of at least one module based on an AI technique. Defined that way, there is evidence of at least four initiatives, all of them driven by judicial agencies of the Autonomous City of Buenos Aires (CABA): those explored by the Criminal and Misdemeanors Court No. 10, the Courthouse No. 13, and the Statistics Office of the Judiciary regarding intelligent anonymization mechanisms of legal documents; and the one driven by the program Prometea, which belongs to the Office of the Public Prosecutor (MPF, for its acronym in Spanish), which in the last years has sought to develop the judicial expert systems through the incorporation of machine learning tools, and it has managed to become a defined and sustained policy in the MPF. From a methodological point of view, it is worth mentioning that the interviews to players of the judicial ecosystem who are driving institutional change initiatives associated with AI use were a critical input of the study about Argentina, which adds an interesting ethnological element to it. We will highlight some key aspects of the study. The first one is that all the cases have been driven by judicial agencies of the sub-national domain, specifically in the CABA, a phenomenon that players themselves think stems from the fact that this district has more resources than the other jurisdictions because it is the wealthiest of the country in terms of GDP per capita. A second outstanding point of the Argentina study is that the institutional innovation processes analyzed are limited to pioneering trajectories with a limited scale and an early degree of development. The only exception is Prometea, a node that has managed to become a defined policy in the Office of the Public Prosecutor (MPF) and which, since 2017, has tried to develop the traditional judiciary expert systems through the incorporation of machine learning tools. Even so, Prometea has also had to resort to alliances, in this case, with the Artificial Intelligence Laboratory of the Law School (IALAB) of the University of Buenos Aires (UBA). On the other hand, the authors state that Argentina does not have defined public policies for the use of AI in justice at a national scale, or a structure of institutional incentives “whose goal is for judicial officials to think they are capable of and responsible for the

use of AI”, for example, by means of awareness and training programs. In this sense, a vital third element of the report is that, in all the cases in which the implementation of AI techniques has been effective, players have resorted to the use of their own human, technical and material resources, and to the creation of formal and informal, intrastate or cross-sectional networks and/ or alliances with other players. The implementations of “intelligent anonymization” encouraged by two of the 31 criminal and misdemeanor courthouses in CABA (No.10 and No.13) were possible from an alliance with another node, whether it is from the judicial sector (Courthouse No. 13 and the Statistics Directorate of the Judicial Council of CABA), or from the private sector with public interest purposes (Courthouse No. 10 and the Technology Cooperative Cambá). Prometea has also resorted to the forging of alliances with other nodes, such as the IALAB of the UBA. A fourth point in the study is the focus on the upward or downward nature of the institutional change dynamics. Thus, ordinary courts have promoted strategies that may be described from a bottom-up dynamic, whereas initiatives such as Prometea describe an institutional change trajectory that is bottom-down by a high-hierarchy agency, in this case the Office of the Deputy Prosecutor of the MPF. The latter explore the figure of institutional entrepreneurs, for example a driver associated with the digital transformation. The fifth outstanding point of the Argentina report refers to the realm of ideas. In line with Isabelle Stengers (2008), who believes that “what ideas we use to think other ideas matters”, Bustos Frati and Gorgone wonder what the previous ideas and beliefs of local experts and authorities regarding justice and AI are, as these ideas may condition the way players think about (and adapt) the principles stated at global governance forums, or in documents of international organizations. On this basis, they identify that, when the time comes to include institutional innovation initiatives based on AI in a broader vision of their ideas about justice, there are at least two frequent characterizations in statements: increased justice and open justice: the former is more axiological, while the latter is technological.

The study about Chile was conducted by Carlos Amunátegui Perelló, Raúl Madrid Ramírez and Matías Aranguiz Villagrán, from the Pontifical Catholic University of Chile. The report points out that the country does not have a specific and complete legislation that regulates Artificial Intelligence, although there is a far-reaching and relevant regulation about digitalization and automation of judicial processes. In particular, the Electronic processing Act 20,886 from 2016 stands out, which established uniform standards for all the jurisdictions regarding the processing of judicial proceedings through an online platform called Virtual Judicial Office. In turn, the

authors mention that the precedent that started the digital transformation process was Memorandum 91/2007 of the Supreme Court, which stipulated the “Rules for Courthouses that process with electronic folders”. In this regard, they stress that the work was influenced by delays in the processing of cases in different jurisdictions, and that it initially met some resistance by the officials of the Judiciary who had been in the institution the longest. Likewise, the study underscores the discussions started in the Future Commission of the Senate, made up by 50 experts in the subject matter from the civil society, to make up working groups on the topic. It also ponders that it is currently discussing a new Civil Proceedings Code, which incorporates important aspects about digitalization and automation, such as the new civil execution model, completely digital and with electronic auctions. However, it points out that the country faces significant challenges when it comes to take the next steps and automate proceedings. In terms of master documents, Chile has the Digital Agenda 2020, which sets a roadmap in the implementation of digital technologies, and it establishes 63 precise compliance indicators. So far, Chile has not developed any official strategy of Artificial Intelligence, but the authors highlight the fact that the Ministry of Science, Technology and Innovation are preparing one following a cross-sectional consultation process. The study also ponders other facts as signs that the country is trying to strengthen the digital transformation of justice, such as the incorporation of the license of Watson Explorer by the Judiciary in 2019, as well as the license of the Watson Knowledge Studio cloud in 2020 for data extraction from web pages (web scraping) to train machine learning methods. Another interesting example, which does not go along the same line of hiring proprietary software, is the agreement signed in 2020 by the Supreme Court and ten universities for the development of a project to update the jurisprudence database. The platform will make it possible, with the use of AI, to access information about Supreme Court sentences easily and quickly. The authors also focus on the reasons why they deem it necessary to develop a sense of urgency around the incorporation of AI in justice. From their perspective, the “demand for judicial services in Chile has not been met” as a result of two main factors: on the one hand, many citizens cannot afford the provision of judicial services because of their high costs, and, on the other hand, there are multiple conflicts that “cannot be tried at court within the current costs structure because taking them to court entails costs that are higher than the possible returns derived from the litigation”. They claim that Artificial Intelligence might help in both fronts. Finally, the report suggests a series of guidelines and principles to lead the prospective incorporation of AI in justice, among which we can highlight the

“responsible transparency”, the traceability of decisions and enabling human intervention whenever necessary, to name a few.

The Colombia report was conducted by Daniel Castaño, researcher of the Externado University. The Colombian case is particularly interesting as it has already moved forward with concrete initiatives to incorporate AI in justice, even high-hierarchy agencies, which sets Colombia apart from Argentina. Likewise, it has established multiple regulatory instruments, albeit most of them non-binding, with regard to digital transformation. We identify the main points of the study. Firstly, the Artificial Intelligence systems that have been incorporated in judicial or procedural matters are identified and described: “Prometea” and “Pretoria” in the Constitutional Court, “Siarelis” in the Superintendence of Corporations in jurisdictional functions, and the “Fiscal Watson” in the Attorney General’s Office. Pretoria is the result of two memorandums (2018 and 2020) signed by the Court and two Argentine universities: the UBA and the Rosario University. The experience of Prometea and the IALAB, already mentioned, has been its conceptual basis. In concrete terms, it is a tool designed to optimize the processing of great volumes of information that the Constitutional Court receives to enable the document management and strengthen the protection selection process, and to improve the quality of the statistical information in the country. Siarelis was implemented to be a guide for users to explore solutions for possible corporate litigations. Finally, Fiscal Watson was implemented to strengthen the Unique System of Criminal Information through the association of cases by means of unstructured databases. Secondly, the report shows that Colombia has a robust judicial framework in terms of digital transformation and social ownership of information and communication technologies (ICTs). Here we find multiple key instruments: i) National Development Plans and, in particular, the Strategic Plan for the Digital Transformation of the Judicial Branch 2021-2025, drafted by the Higher Judiciary Council, which proposes an electronic file model, a network justice system, information management mechanisms, change management methodologies and the use of technology for legal training and citizen services; ii) the Statutory Justice Act and Processing Codes, which encompass legal mandates that favor the implementation and use of ICTs for the management and processing of judicial processes to render access to justice easier and more agile, under the guidance of the Administrative Chamber of the Higher Judiciary Council; and iii) the documents drafted by the National Council of Economic and Social Policy (CONPES), which are non-binding but set important guidelines in terms of digital transformation for public agencies, and even different instruments with ethical

tenets about AI. The Document CONPES 3975 from 2021, which provides a definition of AI, among other aspects, stands out in particular. Thirdly, following the hypothesis that the implementation of Artificial Intelligence in justice administration is not only an opportunity but a need, the study delves deeper into a series of recommendations in terms of public policy. We highlight the following: a) strengthen what the author calls “digital ownership” by justice officials; b) make up a multidisciplinary team with the responsibility for identifying the need they intend to address with the use of AI, and assess the current state in which the task is performed with regard to transparency, promptness, quality, coherence, consistency, efficiency, coordination, inclusion and participation; and c) the implementation of an ethical governance framework of Artificial Intelligence that involves privacy by design, users’ rights and good practices to handle personal information from the architecture. Regarding the last point, Castaño suggests that the algorithms implemented in the administration of justice should comply with the “X-by-design” approach (X= Legality, Privacy or Digital Ethics), according to which the legality, privacy and digital ethics should be part of the technologies, operations and architectures of digital goods and services in a holistic, comprehensive and creative manner. Finally, the study emphasizes that “justice is human and it should remain so”; therefore, it is not about replacing judges with prediction machines. Far from it, the study intends to identify the areas with the greatest potential for the implementation of AI. It focuses on easy, mechanical, repetitive tasks, whose error margin does not pose a risk for the exercise of fundamental rights. Namely, it refers to the automation of processing matters, such as document management, follow-up on notifications and the scope of rulings and other ordinances, to name a few.

The study about Mexico was carried out by the researching team of the Center for Economic Research and Teaching (CIDE, for its acronym in Spanish): María Solange Maqueo Ramírez, Jimena Moreno González, Olivia Andrea Mendoza Enríquez, César Rentería Marín, and the assistance of Claudio Andrés Ambroglini Gómez. The report stands out because of its comprehensive approach to the institutional complexity of the judicial domain in a country with a federal political regime. Thus, even though there are no cases of AI use in the Mexican justice, at none of its levels, the report lays the regulatory, doctrinal and institutional foundations to project its future integration in the judicial domain. Along this line, the study identifies different factors that hinder the best preparedness to incorporate AI. Mainly, the lack of a homogeneous model of the Judiciary along the federative entities, which is verified by a significant diversity in terms of

size and design of courthouses, chambers and administrative; and the great asymmetry between the Judiciaries in their digital updating, especially regarding the production, storage and processing of data that are relevant for decision-making in the administration of justice. The differences are even verified in the ways each jurisdiction addresses the process of digital transformation, as is the case of the digital signature. The federal Judiciary has digital signatures (FIREL), but the state Judiciaries have taken different steps; some of them have developed their own digital signatures, whereas others are adopting the FIREL or similar as their standard. In turn, in many cases, the functionality of the electronic file management systems available locally is limited to capturing and consulting digital documents. A third contrast has to do with the availability of an administrative unit of legal statistics that is in charge of the production and processing of data and metadata of electronic files. Indeed, in this regard, the authors consider that the Global System of File Follow-Up of the Federal Judiciary constitutes a best practiceFrom the regulatory standpoint, even though Mexico does not have specific legislation about AI, let alone a legal framework that addresses the current administrative and civil responsibilities derived from AI use specifically, the report considers that the present constitutional and legal order encompasses enough elements to face, in a general way, the damages that AI systems may cause to people. In particular, with regard to the regulation about legal data management, the document identifies a trend in the last five years, since the passing of the general Transparency and Access to Public Information Act from 2016, to make several changes. Additionally, the mapping of multiple specific documents around this subject matter, issued by different players of the Mexican judicial system, stands out. Regarding strategies, the study points out that there is no national strategy that leads players of the digital ecosystem for an action plan or that gives guidelines for the development of AI in the public sector. In the national digital policy, outlined by the National Development Plan and its respective programs, no possible advancement is explicitly projected that are related to AI. Now then, Mexico has had an AI Strategy since 2018, but the authors stress that it was published in the last days of the Enrique Peña Nieto administration, so its implementation has been rather limited, while there is no evidence that the current government of Andrés Manuel López Obrador intends to follow up on it. Another outstanding aspect of the study is the identification of international sources of the Mexican Law on the subject. The most prominent ones are the alignment with the European Union (data protection in line with the standards of the European Regulation), the Council of Europe (through the endorsement of Convention 108 and the so-called Convention 108 Plus) and the

Organization for Economic Co-operation and Development (OECD) (through the AI Principles), among other forums. Likewise, its participation in the Digital 9 Consortium (together with Uruguay) and the Ibero-American Data Protection Network (RIPD) where Mexico has played a central role in the regional coordination is highlighted. Finally, a valuable aspect is the notion that the context of the COVID-19 pandemic has stirred up the interest of lawmakers to create regulatory frameworks that refer specifically to this topic, which is expressed in the existence of four legislative initiatives, which are currently awaiting a decision. On the other hand, even though some precedents have been identified in terms of legal expert systems, none of them are currently in place.

Finally, the study about Uruguay was conducted by Julio Lens, Sandra Segredo and Fernando Vargas, from the Catholic University of Uruguay Dámaso Antonio Larrañaga. The pivotal point of this report is the proposition that Uruguay seems to be undergoing a “twospeed technology reception or progress”. This is so because there are two different scenarios depending on the state environment. Thus, government agencies appear to be “reasonably receptive to new technologies and moderately agile in their detection and incorporation”, while in other state areas, especially in the Justice, such progress is not taking place, or else “when it does, it is in a particularly cautious manner, significantly slower”. Indeed, in terms of informational development and digital transformation of the Executive Branch, Uruguay is one of the countries with greater trajectory at a regional level. The study explores the participation of the country in Digital 9 (Mexico is also part of the forum) and its realization in the Digital Nations Charter, as well as the efforts of the Agency for the Development of the Electronic Management Government and the Information and Knowledge Society” (AGESIC). It also addresses the digital plans deployed since 2006, renewed today around the Uruguay Digital Agenda, which have been translated into measures such as the implementation of the plans Ceibal and Ibirapitá, a unique integrating portal that enables online state processing, the digital file in the Public Administration, electronic communication and notifications, digital medical histories and the use of blockchain in public records, among other development (Aoudie, Ast, Deffains, 2021). At the same time, the study seeks to describe the institutional precedents in the judicial domain and, based on the above, it is a descriptive exercise that yields clearly asymmetrical results. For example, at a Judiciary level, there is not any comprehensive electronic judicial file that makes it possible to implement judicial processes in a totally electronic format. In fact, here the precedents that affect technology governance in the

judicial domain are rather landmarks driven by the State as a whole, such as the Public Information Access Act (2008) or the electronic file (2009), which “created the framework necessary to install a new model of management for courthouses and supporting offices”. The study makes an effort to identify and map the digital capacities developed in the judicial field, such as the National Jurisprudence Database (in place since 2008) or the Multisubject Courthouse Management System (an IT application to follow up on processing at courthouses and courtrooms, as well as the systematization, preservation and consultation of all the proceedings). Now then, the most suitable element singled out by the study is that, recently, the Jurisprudence Department of the Judiciary started a pilot trial with the Engineering School of the University of the Republic for the anonymization of court decisions. So far, there has been no definitive report about the results produced by the enterprise, but it is an encouraging initiative. Meanwhile, from the regulatory standpoint, the report makes it clear that there is no national regulation concerning the use of Artificial Intelligence, or for government matters, and even less an explicit regulatory reference about its implementation for the processes carried out by the Judiciary. There is, however, a National Strategy of Artificial Intelligence for the Digital Government, published in 2019, which lays down the principles for the development, use and implementation of Artificial Intelligence systems in the public sector in the country. The authors also highlight, as a “guiding document”, the text “Questions for the assessment of the algorithmic impact study”, prepared by the Information Technologies Area of the AGESIC in 2020, as it is a “guide that applies to the system projects that use automated learning for decision-making”.

Mapping of similarities and differences between national cases

mework has allowed us to identify some common aspects and certain contrasts that are worth mentioning (see Figure 1) to map the state of the discussion, and of the preparedness in terms of ideas and imaginary scenarios, (pre) concepts, policies and legislation on this subject. Figure 1 – Summary of national reports National digital strategies Argentina, Chile, Mexico and Uruguay have government strategies. Colombia has a digital strategy for the Judiciary. AI government strategies Argentina, Chile, Colombia, Mexico and Uruguay. With different degrees of implementation and continuity in the policies outlined. AI strategies in Justice None of the cases analyzed, even though in Colombia the digital strategy for the Judiciary mentions data analytics

and provides a definition of AI. Specific regulations for AI None of the cases analyzed, even though Chile has a specific provision in the DEPA; in Colombia there are doctrinal non-binding documents. Effective initiatives about AI in Justice Argentina and Colombia Use of expert systems in Justice Argentina, Colombia and Mexico (they are no longer in place in Mexico) Institutional entrepreneurs and upward institutional change dynamics through the use of AI Argentina Downward institutional change dynamics through the use of AI Argentina and Colombia The first thing to point out is that, in all the cases, a kind of digital transformation process of twospeed public agencies is verified, where the dynamism observed at government level is not verified in the Judiciaries. This is so because of different factors that each national report addresses differently, but, by and large, they agree on the opacity of the judicial practice, the lower citizen control and judges' reluctance to change. Strategic alliances for the implementation of AI in the Judiciary Argentina (all the alliances are informal, except for those generated by the IALAB) and Colombia (formal agreements with Argentine universities). The Courts of Uruguay and Chile have reached agreements with universities that may translate into AI-based initiatives. Speeding up digitalization processes after the pandemic In all the cases. In fact, in Colombia the digital strategy for the Judiciary was drafted within the framework of the pandemic. Potential applicability of AI in Justice All the reports agree on the fact that the automation of some mechanical and routine tasks, where it is not necessary to interpret a complex context, is an opportunity, as it would make it possible to take better advantage of the resources of the judicial system. With regard to the evidence of effectively existing initiatives about the use of AI, it is only found in the cases of Argentina and Colombia. In the case of Argentina, they are initiatives driven by judicial bodies at sub-national level, whereas in Colombia we identify initiatives in the highest-hierarchy judicial agencies at national level, such as the Constitutional Court and the Attorney General's Office. A key similarity, however, is the fact that initiatives in both countries have been possible as a result of the convergence or alliance between different players. That is to say, cross-sectional alliances appear as a vector that favors the digital transformation of justice based on AI use. That has translated into initiatives for the effective use of AI both in Argentina and in Colombia, but it can also be verified that the two most interesting initiatives around the use of AI in the Chilean and Uruguayan Justice are related to agreements reached with universities. In the case of Argentina, the role of institutional entrepreneurs, who drive processes of upward institutional change of their own accord, stands out. Also in institutional terms, which have necessarily been manifested in AI

governance in the judicial domain, the fact that the two countries with federal regimes (Mexico and Argentina) showcase a special complexity is highlighted; such complexity is brought about by the co-existence of local and federal courthouses. Now then, if the report about Mexico keeps the focus at the federal level, the report about Argentina is forced to explore the sub-national dynamics, as it is there where we see the greatest dynamism in that subject matter. In the other three countries, we find unitary regimes, which favor the unification of standards and processes, even when, naturally, this does not entail a preference in favor of incorporating AI. On the other hand, none of the countries analyzed has a specific regulation on AI, or national strategies for AI in the judicial area. The most widespread practice is the existence of national digital strategies, albeit restricted to the government domain. The only country that has incorporated a specific digital strategy for the judicial realm is Colombia, in 2020, from the challenges brought about by the pandemic. Even though there are documents as national AI strategies in Argentina, Mexico and Colombia, none of them addresses aspects related to the judicial domain specifically. Additionally, there is only evidence that it is implemented in Colombia. Indeed, both in Argentina and in Mexico the strategy was published in the last days of the governments that promoted their processes of cross-sectional consultation and drafting (Mauricio Macri and Enrique Peña Nieto, respectively), so their implementation has not been materialized in the administrations of the new national governments yet, and they serve as reference documents rather than current public policies (Aguerre, 2020). As for national digital strategies, they can be found in the five countries under analysis, although some of them have greater precedents in the subjects (such as Uruguay, which has submitted several strategies since 2006). Another common point in some of the works is the idea that the use of AI in the judicial domain is an opportunity (and not just a risk) given the expectation that the automation of some tasks and areas (different from those where AI is not applicable) might make it possible to direct the human talent to the critical aspects of the judicial system. Here, the reports of Argentina, Mexico, Colombia and Uruguay concur. Moreover, there is some consensus in that the most complex tasks, like the ones that entail the interpretation of a context or the approach to a case for which there is no clear legislation, should not be automated, in line with the uses recommended by the CEPEJ of the Council of Europe (COE, 2019). Finally, there is agreement that human beings should always be responsible for the decisions made; therefore, they suggest that judges will see their roles modified but not replaced. In turn, a significant similarity is that all the reports see the Covid-19 pandemic as a contextual opportunity,

as it has created a sense of urgency among political decision-makers. Indeed, all the reports agree on the fact that the processes of judicial digitalization have been expedited since the beginning of the pandemic, even when different journeys may be observed. By way of conclusion, it is relevant to point out that we find the most relevant progress in the discussion, objectives and aims regarding the implementation of this technology in the judicial domain in those fields where there is work based on dialogue with other sectors and players. More important than the level of digitalization reached by a country or State is the capacity to integrate levels of discussion that consider the macro, meso and micro levels as a key dimension to coordinate the objectives of this "automation", as well as the different perspectives associated. Based on the experience of this research, the Judiciary needs to address the inclusion of all the advanced AI technologies in their systems considering the broader narratives taking place around AI use in other sectors, in other regions and with other players, as the values that should be protected are fundamental safeguards of democratic societies.

36. National Artificial Intelligence Strategy

Artificial Intelligence

Classic definition Due to the accelerated development of Artificial Intelligence (AI), its definition has been transformed to the point where some consider that AI is everything that has not yet been invented. In order to build its understanding, we will start by mentioning that the term was coined by Professor John McCarthy in 1955. For him, AI was "the science and engineering of making intelligent machines. " Around the same time, mathematician Alan Turing also proposed a test to evaluate whether a machine could impersonate a human in its interaction with a real human.

Definition for the Strategy

- In its early days, AI simulated human behavior through rules. However, more recent methods focus on making machines learn from data and their interaction with the outside world automatically. Currently, the OECD proposes the following definition: "An AI system is a

machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy" . • To complement these two definitions, we would also like to mention that AI is influenced by and influences areas such as neuroscience, linguistics, robotics, among others.

Machine Learning (Automatic or machine learning)

- Set of Artificial Intelligence models trained to learn from data in order to predict outcomes or make decisions without being explicitly programmed to do so. • Some of the best known types of approaches are: supervised, unsupervised, reinforcement and association rule learning, etc. • Some of the best known algorithms are: artificial neural networks, decision trees, support vector machines, Bayesian networks, genetic algorithms, deep learning and others.

External Context

The data economy and Artificial Intelligence applications are being adopted around the world at such a speed that, for the last 5 years, several nations and multilateral organizations have taken a more conscious stance on the need to promote and govern this process through the design and implementation of National Plans, Strategies and Agendas. Evidence on international trends (*): • In recent years many countries around the world have invested in Artificial Intelligence led by the US and China. • Use cases in almost every industry are accelerating. • Since 2015, the number of AI-related jobs in the private sector has doubled • Academic offerings for specialization, Master's and Doctorate degrees have increased significantly. • AI publications are increasing significantly. • AI technology or consultancy offerings are growing at an accelerated pace • Investments in technological infrastructure to support increased adoption and use. • The number of PhD graduates with AI and Machine Learning has doubled. • Multilateral organizations and Nations are defining Ethical Frameworks and regulations specifically dealing with AI. • Private investment in AI projects is increasing. <https://hai.stanford.edu/research/ai-index-2021> Photo: Pixabay External Context In our region, Brazil, Mexico, Colombia, Argentina, Chile, and Uruguay have already made progress in this area, with various documented initiatives. In all cases, the documents reveal expressions of interest in the direction that governments should take with respect to these

phenomena, and many of them are expressed in public policies but with little empirical support or technical basis. However, the cases of Colombia and Brazil have made significant progress in the implementation of their plans, which include the creation of institutions or agencies specially designed to manage strategies, coordinate efforts, and channel public financing. In all cases there was a great effort of socialization, dissemination, and involvement with the different stakeholders. The main fronts addressed for the adoption and use of IA were:

- Closing specialized talent gaps.
- Ensure the provision of actionable data and operations infrastructure.
- Promotion of adoption-oriented investments in sectors such as small and medium business and startups.
- Definition of Ethical Frameworks and specific regulation.

External Analysis: Use Cases of Local Government

Improved citizen service

- Several municipalities in Norway use chatbots to speed up customer service for simple requests from citizens.
- The municipality of Copenhagen uses NLP technique to analyze and sort through incoming emails from citizens.
- Decision support
- The municipality of Espoo (Finland) developed a pilot project to plan the budget allocated for public health services.
- The municipality of Trondheim (Norway) developed an administrative case recommendation system with the aim of speeding up decision making given a similar request.

External Context: Use Cases in Justice:

Robot Judge

- China has been employing artificial intelligence in the courtroom since 2017. A robot judge is used to hear specific cases such as trade disputes, e-commerce liability claims, and copyright infringement. To date, more than 3 million cases have been handled by a robot judge in China.
- System to predict future criminal recidivism
- HART is a system developed by the Durham Police and Cambridge University researchers, using training data from 104,000 people who had been arrested for five years.
- The system uses variables that focus on the suspect's crime history, as well as age, gender, and geographic area, to classify an offender as low, medium, or high risk of committing new serious crimes during the period of next two years.
- HART was developed with the aim of reducing the number of people incarcerated, and of being susceptible to other forms of intervention that would be as or even more effective in reducing the risk of recidivism

External Context: Use Case in Citizen Security

Use Cases in Public Transport • Within public transportation, include cameras and AI-based systems to be able to detect people fights, follow people, agitators, or suspicious behaviour or any activity that could be a danger to women. It could also detect areas that do not have much light. • These systems could alert operators to react quickly. • AI System to predict crimes in specific places and times • PredPol, a Californian company that grew out of a project between UCLA and Los Angeles Police Department, defines "predictive surveillance as the practice of identifying the times and places where specific crimes are most likely to occur, and then patrolling those areas to prevent for those crimes to occur. • PredPol uses the historical data of a client's police department from a period of two to five years to train a machine learning algorithm, which is subsequently updated on a daily basis. Only three data points are used: crime type, location, and date/time. • According to PredPol, demographic, ethnic or socioeconomic information is never used. This eliminates the possibility of privacy or civil rights violations seen with other predictive or intelligence-based policing models.

External Context: Use Case in Education

AI adapted to the needs of each student • From school to college, AI could individualize the learning needs of each student. • The system could respond to the needs of the student, placing greater emphasis on certain topics, repeating things until the student masters it. • In general, helping the student to learn at her own pace, whatever this could be. • AI could give feedback to students and teachers • AI could give feedback to teachers and students about the results of the course itself. • Some AI systems are used to monitor student progress and alert teachers when there might be a problem with student performance. • These systems could give the support to the students who need it, and to the teacher, to find the areas where he can improve the instructions to the student so that he does not fail with the subject of the course.

External Analysis: Use Cases in Agriculture

- Digital production and agronomy Initiatives dedicated to crop monitoring through sensors and images from drones and satellites).
- Crop planning and management Managing crops, irrigation and workers based on data and predictive analytics
- Market access and financing Crop and cost forecasting, insurance, exchange applications and leftover recovery betterfoodventures
- 50% water and energy savings
- Reduction of losses on expensive fertilizers due to overwatering

Use Cases in Fisheries and Aquaculture

- Reduction of operating costs Reduce the cost of farm maintenance, optimization of inputs and resources.
- Observe Technologies and Umitron Cell offer AI-based technologies for fish feeding frequency pattern detection
- Managing fish health
- Norway's Seafood Innovation Cluster developed an ML system to predict possible parasite outbreaks in fish farms

Use Cases in Forest Protection:

- Forest resource management Forest monitoring through satellite images.
- SilviaTerra and 20tree.ai use deep learning algorithms to generate highly accurate forest maps and extract attributes such as tree species, size and diameter, etc.
- Deforestation detection Prediction of illegal logging activities.
- Rainforest Connection and Outland Analytics use sensors and deep learning models to detect suspicious sounds such as chainsaws or heavy machinery.

Use Cases in Telecommunications:

- Equipment and Network Optimization Traffic monitoring to detect problems in the communications equipment or to self-manage the traffic
- Aria Networks and Avaneseus use ML algorithms to detect traffic anomalies and anticipate potential network problems.
- Improved customer service Implementation of automatic question and answer system that can reduce the response time to a network problem.
- Success.ai or Vodafone use chatbots systems to solve the most frequently asked questions regarding a network problem

Use Cases for the Mining Industry:

- Process improvement • Analysis of machinery operation to determine maximum performance.
- Freeport-McMoran achieved a 5% increase in copper production by using ML to increase mill throughput at its Arizona mines.
- Improved scanning • New methods for the analysis of images and other sources of information to find mineral deposits.
- Datarock uses deep learning and computer vision to extract geological information from geospatial imagery.
- Minerva Intelligence uses semantic web technologies to harmonize databases with geological information for easy analysis.

Use Cases in Energy:

- Energy demand forecasting • In the UK, National Grid and DeepMind developed an ML algorithm to predict load in the short term, reducing energy consumption by 10%.
- Optimize power generation • GE Renewable Energy uses software that monitors and optimizes the operation of its wind turbines, increasing energy production by 20%

Ethics in AI:

- Biases and discrimination • Data and algorithm evaluation • Multidisciplinary evaluation • Job automation
- Difficulty of measurement - Informality • Promote digitization jobs • Promote training
- Guidelines, worldwide guidelines and proposals • Identify negative stereotypes and usefulness of personalization (access to university, access to credits, medicine and others).
- Promote use in vulnerable population problems, for health and logistic applications • Constant updating

Internal Analysis:

In Peru, some progress has been observed in the development and adoption of Artificial Intelligence, mainly from large economic groups and a small sector of academia; while the public sector has made very few attempts . These use cases have been centralized mainly in Lima. These are some of the situations encountered :

- The country's main economic groups are carrying out or have carried out AI projects .
- There are AI projects in the private sector, most of those projects are Machine Learning, few in Deep Learning .
- It has been observed that the number of

publications in Scopus Indexed Journals with an AI component has quadrupled in the last 5 years in the country, mainly due to incentives provided by concytec, innovate and other public funds . • The main universities that publish on AI topics are in Lima, followed by Arequipa. • The areas of publication with some AI component are mainly Computer Science, Engineering and Medicine with 50% of the total number of publications . • Currently many Peruvian universities have master's programs in data science, in addition to master's and doctoral programs with AI courses .

Main organizations using AI in Peru:

- Main economic groups: Breca Group (Brescia family), Intercorp (Carlos Rodriguez Pastor), Belcorp (Eduardo Belmont). • Companies / Large Corporations: Banks, Telcos, Retail, Insurance, Mining / Fishing, Manufacturing, etc. • Startups: Quantum Talent, Emptor, Fitness Pass, Xertica, Teckton Labs, Latin Fintech, Chazki, qAIra, SpaceAg, etc. • Consulting Firms: McKinsey, EY, MS, DMC, Everis, Accenture, Globant • Technologies (and their partners): Google, Microsoft, AWS, IBM, etc. There is evidence that in some cases everyone uses Machine Learning, some NLP, little Deep Learning.

Labor market situation in Peru Photo:

Pixabay The recruitment of professionals with knowledge of AI (or related subjects) in different sectors: • Private Sector: Corporations, Companies, Startups, Consulting Firms, Technological Firms Salary ranges: 95% of the salaries are between S/. 2,000 - S/. 20,000 (min: S/. 2,000 1st quart: S/. 4,000 mean: S/. 6,000, median: S/. 7,500, 3rd quart: S/. 12,500) • Academia: As research professors in Engineering, Computer Science, Diplomas, Business Schools. • Government: As hired staff in public institutions like Government and Digital Transformation Secretariat, SUNAT, MEF, Minedu, Comptroller General of the Republic, Judiciary and others. • Otros sectores: Multilaterales (BID, Banco Mundial, CAF), ONGs / Think tanks (Innovations for Poverty Action - IPA / JPAL, etc.)

State of the local academy in terms of training of professionals with AI knowledge:

Undergraduate Studies: • Traditional Computer/Systems/Computer Engineering, Computer Science careers that have been updating their curricula. • Other related majors - Economics, Mathematics, Physics have begun to focus on foundational support or AI courses. Postgraduate studies: • PUCP - Doctorate Program in Computer Science • UNI - M.S. and Ph.D. programs in systems engineering with AI courses, M.S. in Computer Science. • URP - Masters in Data Science • UPC - Master in Data Science • UNSA - Master in Computer Science • UP - MBA in Business Analytics, Diploma in Analytics • UDEP - Diploma in Data Science (with Pompeu Fabra University) • UTEC - Masters in Computer Science with specializations, Diploma in Data Science for Business • Several institutions that teach Data Science / Analytics / AI courses: DMC, UNI, Colectivo23, BetaHouse, Kurios, Instituto de Analítica Avanzada (Breca), Digital House, Crehana, etc

Results of the Peruvian Public Sector AI Status Survey (January - 2021) Question:

In which use cases do you apply AI or Data Science in your organization? • Prediction of criminal acts • Facial recognition for assistance taking • Use cases in agriculture such as identification of crop types using satellite imagery • Natural gas and electricity monitoring using AI and regulation automation with Big Data • AI In Social Programs • Control of users of basic services (electricity and water) • Tax applications • Risk management • Virtual assistants • Customs risk management

Applications with AI Components Developed During the COVID Pandemic in Peru:

"Peru in your hands" developed in conjunction with private companies with expertise in mobile applications, artificial intelligence and data analytics: Team: (UNI), (UP), (UTEC), MIT, Stanford University, INSEAD Paris, Universitat Pompeu Fabra & Barcelona GSE & IPEG & CEPR, Tekton Labs, Kambista, Sapia, Mr. Burns, Media Labs, AmigoCloud, Alicorp, the Korean Ministry of Interior, the Andean Development Community and the Inter-American Development Bank. Project "COVID: Dynamic Virus Control". Winner of the "Special Projects: Response to COVID-19" contest organized by Concytec. Proposal: dynamic control of virus infection in the Peruvian population with the use of artificial intelligence. Two processes: First: detection of potential COVID-19 infectees. Second: pre-diagnosis of the virus. Intelligent Integrated System for recording, reporting, alerting, monitoring and assisting people symptomatic of COVID-19

(SIAMA)." "SIAMA", a system to inform, alert, monitor and assist COVID-19 symptomatic people in Latin America, when delivering data and information does so through its App or web platform, generating automated alerts. The innovative system has a voice assistant, which, by means of periodic questions, will be informing the nearest healthcare personnel of the infected inhabitant. What technologies do you use? SIAMA uses Deep Learning (AI technique) that integrates a chatbot based on natural language. It provides patients with a better order in the level of urgency, in addition to the integration of the voice assistant (Google home or Alexa), which allows the interaction of symptomatic people, through voice commands; Smart Security Office" platform to control Covid-19 Company: MDP Consulting Winning project of the InnovaCovid-19 Challenge, of the Innóvate Perú Program, will receive a co-financing of up to S/. 450,000 Manages and controls the access of employees and visitors to work centers, through identity validation, detection of mask use, body temperature measurement and rapid triage for the prevention of Covid-19. Technology: Facial recognition, temperature sensor to monitor the worker's health status and detect if he/she is wearing a mask. It allows for periodic triage of personnel through a chatbot

SWOT Analysis:

- Strengths • The country has scientific, mathematical and engineering talent.
- Large economic groups interested and investing in AI • Government interested and promoting through PCM, CONCYTEC, FONDECYT, Innóvate and others.
- There is a network of contacts with universities and top companies.
- Programs such as Beca Presidente • Peruvian universities researching these topics Opportunities
- Online Training (MOOCs)
- Growing labor demand in the private sector, public sector, and academia for these AI positions (including from major companies such as Amazon, Google, etc.).
- Growing interest and funding from Multilateral organizations
- Peruvian Diaspora in top places
- More Peruvians studying and researching IA in the world
- The needs of attention to SMEs Weaknesses
- Lack of high-level Educational Opportunities
- Very little development of advanced AI (Deep Learning) in the private sector.
- Incipient government on IA issues
- Data situation at the country level, and at the level of organizations (private, NGOs, etc.).
- Soft skills, English, and talent research capabilities
- Salary incentives for academic researchers in AI are low.

Threats • Brain Drain: growing labor demand in the private sector from major Amazon, Google, etc.

- Other countries more AI advanced than us
- Political instability
- Few institutions in Peru promote AI development
- External sector attracts "all" talent
- Peruvian society may have some fears about the use of AI as its use is further promoted.

Purpose of the National Strategy for Artificial Intelligence:

For a country like Peru, the adoption of new technologies such as Artificial Intelligence represents a historically unprecedented opportunity and at the same time a great threat to be left behind in the global socio-economic development and to deepen our shortcomings. This historical singularity presents an enormous responsibility for all members of our Society, not only the State, but also for the Private Sector, the Academia and the Civil Society. Then it becomes imperative the will to govern this adoption process with a National Strategy for Artificial Intelligence (ENIA) that facilitates the conditions to take advantage of these opportunities and mitigate the risks derived from this process. It is clear that the functioning of Society has been digitized on almost all fronts, driving a new economy whose main asset is Data. Creating value with these assets is an imperative not only for the private sector, but also for the public sector, which must also meet the new demands of a digital citizenship. In this context, the most advanced nations have deployed efforts during the last 4 years to express their intentions to govern a process that turns AI into a tool for socio-economic development and that is not a consequence of chance. This has been the main purpose for which various manifestations have been proposed, ranging from Strategies, Plans and National Agendas that in their first stages contain diagnoses of the starting point and aspirations for the end of this decade, through public policy recommendations, strengthening of some sectors and prioritization of public-private initiatives to ensure the first steps towards a National Artificial Intelligence Strategy. The following National Artificial Intelligence Strategy (ENIA) is proposed for the period 2021-2026, which can be updated every 2 years according to new technological advances and the situation of the country and the world. Peru is recognized as a Latin American leader in research, development, innovation, deployment, use, adoption of AI, and in its ethical and responsible use in the production of public and private goods and services. These efforts aim to accelerate national development and promote digital inclusion while ensuring the reduction of social gaps. Implement the National AI Strategy, promoting the development of human talent, infrastructure and the production of goods and services based on AI for the benefit of a more inclusive and multicultural Peruvian society that takes advantage of Industry 4.0 technologies in favor of sustainable development

Strategic objectives:

E1 - Training and Talent Attraction

SO.1.1. To position Peru as a country that enhances its human talent at all educational levels for the research, development and uses of Artificial Intelligence in the country. SO.1.2. To Lead regional research, scientific publication and patent publication in AI in key sectors of the country. SO.1.3. To be an attractive country for AI research and development. SO.1.4. Reduce the gap of participation of women and minorities in AI training programs.

SO.1.1. To position Peru as a country that enhances its human talent in Artificial Intelligence at all educational levels for research, development and use of AI in the country. • A.1.1.6. Create master's and doctorate programs in AI / ML in universities in the country with the formation of CORE programs on Artificial Intelligence. • A.1.1.7. Create scholarship programs for Doctoral students from IA and ML programs. • A.1.1.8. Promote massive up-skilling and re-skilling programs in digital skills, and in ethical issues of data processing. • A.1.1.9. Create courses or diplomas for the training of talent in the installation, administration and infrastructure support of high-performance supercomputers. • A.1.1.10. Create courses in parallel computing, signal processing and other courses related to high performance computing within the undergraduate and graduate programs of computer science, ing. software, ing. systems, ing. Computer science and other related.

SO.2.1. To Lead regional research, scientific publication and patent publication in AI in key sectors of the country. • A.1.2.1. Increase the number of publications in IA / ML according to RENACYT standards (in international conferences). • A.1.2.2. Encourage the creation of patents resulting from competitive funds and / or alliances with the private sector. • A.1.2.3. Facilitate international AI conferences that are held in different regions of the country on the following topics: AI applications in agro-industry, health, mining, forestry protection, energy, fishing and government, and in other sectors that the National Center for Innovation and Artificial Intelligence recommends it. • A.1.2.4. Facilitate visiting professor programs between local and foreign universities (including professors who teach remotely at provincial universities)

SO.1.3. To be an attractive country for AI research and development.

- A.1.3.1. Create programs to attract Peruvian or foreign talent with Doctorate degrees in AI/ML from the academy, private or public sector, implementing incentive programs; for example, tax incentives for the private sector, or research funds for academia.
- A.1.3.2. Create programs to retain talent such as PhD in AI/ML graduated in Peru or abroad, in research centers and universities in the country. These policies could consider better salaries, education and training opportunities, job stability and appointments, infrastructure improvements and others.
- A.1.3.3. Facilitate the National Center for Innovation and Artificial Intelligence to repatriate national or foreign talent for its research and development projects in AI.
- A.1.3.4. Facilitate public Peruvian universities to enter into agreements with top universities in Europe or USA in their AI/ ML master and doctorate programs.
- A.1.3.5. Facilitate the processes of validation of professional Master's or Doctorate degrees carried out abroad

SO.1.4. Reduce the participation gap of women and minorities in IA CORE training programs.

- A.1.4.1. Promote the largest number of female students in undergraduate and graduate CORE Artificial Intelligence (AI) training programs at Peruvian universities.
- A.1.4.2. Promote the decentralization of AI/ML master's and doctorate programs to different regions of the country, promoting in regions where there is great potential as part of the country's key sectors, such as agro-industry, cattle raising , mining, forestry protection, health, fishing.
- A.1.4.3. Create programming teaching programs for the unemployed and adults through regional or municipal centers in the different regions of the country.

E2 - Economic model

SO.2.1. Leading the research and development of Artificial Intelligence at the regional level.

SO.2.2. Promote in public bodies, the incorporation of artificial intelligence in their operation and services to citizens.

SO.2.3. Promote the integration of AI in the value chain to promote business development in the country's key economic sectors.

SO.2.4. Minimize the effect of job displacement due to the adoption of AI

SO.2.1. To lead the research and development of Artificial Intelligence at the regional level.

- A.2.1.1. Increase research funds in AI through the different programs financed with public resources.
- A.2.1.2. Improve the evaluation of research and innovation projects with public funds

with AI components by convening more experts in the area. • A.2.1.3. The winners of projects financed with public resources must upload their data and codes (models) to the open data platform of the government of Peru to be shared. • A.2.1.4. Promote scientific dissemination initiatives in coordination with the different research centers and universities. • A.2.1.5. Monitor the index development and adoption of AI (*) and act where necessary to be in the first place at the regional level.

SO.2.2. Promote in public bodies, the incorporation of artificial intelligence in their operation and services to citizens. • A.2.2.1 Develop online courses for public officials in the adoption, use and benefits of Artificial Intelligence. • A.2.2.2. Promote the use of technologies such as chatbots or virtual assistants in public administration, prioritizing libraries and open source software. • A.2.2.3. Through the National Center for Innovation and Artificial Intelligence, prioritize the development of use cases where Artificial Intelligence can generate concrete solutions such as those proposed in various investigations aligned to the United Nations 2030 sustainable development goals, such as the elimination of poverty, zero hunger, quality education, clean and accessible energy, clean water and sustainable cities, good health, better qualified jobs, the reduction of social gaps and others. • A.2.2.4 Promote citizen participation in IAckathon, Datathon or Hackathon events organized by public bodies to create new services or improvements in citizen care, or to solve some problems in our country such as the detection of corruption in the public sector. • A.2.2.5. Promote the use of open source technologies in the public sector for the development of use cases based on Artificial Intelligence.

SO.2.3. Promote the integration of AI in the value chain to promote business development in the country's key economic sectors. • A.2.3.1. Create public funds and increase those that currently exist, for startups and basic or applied research programs that have some AI component, prioritizing the applications and areas of knowledge that the National Center for Innovation and Artificial Intelligence considers key. • A.2.3.2. Promote collaboration between academia and industry (micro, small and medium enterprises), also integrating CITEs from all regions of the country. • A.2.3.3. Promote initiatives to generate local use cases based on Artificial Intelligence with open data from the private, public and academic sectors of the country. • A.2.3.4. Promote a labelling industry through financial incentives for new ventures as part of publicly funded startup financing programs. • A.2.3.5. Prioritize and promote the adoption of AI in strategic economic sectors for the country, where we have comparative and competitive advantages.

SO.2.3. Promote the integration of AI in the value chain to promote business development in the country's key economic sectors. • A.2.3.6. Incorporate in each project financed with public resources the use of the infrastructure of the existing local data centers in the country, or in cloud platforms of private providers but where their technological infrastructure is installed in the country. • A.2.3.7. Promote research on Natural Language Processing algorithms in Spanish and other native languages. • A.2.3.8. Promote incentives for the private sector to publish open data on the country's national open data platform. • A.2.3.9. Promote regulatory sandboxes for key economic sectors of the country as environments for business development (small, medium, large companies, as well as startups) with few regulations in use cases with AI components, promoting their development and innovation in the country.

SO 2.4. Minimize the effect of job displacement due to the adoption of AI. • A.2.4.1. Create an observatory for the labor market to obtain timely evidence on the effects of AI. • A.2.4.2. Promote policies for upskilling / reskilling of the population that will be affected by the adoption of AI

E3 - Technological infrastructure

SO.3.1. Improve local infrastructure for AI research and development. SO.3.2. Promote the integration of local infrastructure. SO.3.1. Improve local infrastructure for AI research and development • A.3.1.1. Consolidate the National Fiber Optic Backbone Network project, in addition to installing a high-speed fiber optic connection between research centers. • A.3.1.2. Promote the Implementation of 5G technology by spreading its benefits. • A.3.1.3. Create a National Center for High Performance Computing for research from academia, public institutions and industry in different areas of knowledge. • A.3.1.4. Promote the public-private partnership for the installation of the infrastructure of data centers in the cloud in the country with services for the academy, public sector, private sector and the citizenship in general of the country. • A.3.1.5. The government will promote contracting for the public sector, academia, local small and medium businesses, cloud services but with a provider that has local infrastructure and at a low cost. The government could guarantee a minimum monthly demand to be able to incentivize the installation of local cloud infrastructure in the country. • A.3.1.6. Increase investment in educational infrastructure for the training of new talents in AI that includes implementation of IA / ML laboratories in the different universities of the country

SO.3.2. Promote the integration of local infrastructure. • A.3.2.1. Through the National Center for High Performance Computing to govern all high performance computing centers and those created in recent years with public funding. The supercomputers have to be connected with a high speed optical fiber. • A.3.2.2. Rationalize the use of public funds for the acquisition of infrastructure equipment for research by facilitating the use of high-performance computing centers or local data centers once implemented. • A.3.2.3. Train local talent in the installation and administration of high performance centers in different regions of the country. • A.3.2.4. Create incentives to promote the use of high performance computing centers and local data centers in the country in research and development programs with public fund

E4 - Data

SO.4.1. To be a regional leader in the publication of open data. SO.4.2. To be a regional leader in the publication of data on biodiversity, native languages and other minorities in the country

SO.4.1. To be a regional leader in the publication of open data. • A.4.1.1. Create a rewards or penalties program for public and/or private organizations to publish high-quality open data in an open, reusable format on the country's government's open data platform (datosaberto.gob.pe). • A.4.1.2. In the datosabiertos.gob.pe platform include a module for communities to store codes and AI models based on the datasets published. • A.4.1.3. Create APIs to extract information from the single platform gob.pe and datosabiertos.gob.pe. • A.4.1.4. Monitor existing open data barometers (*) and act if necessary to keep the country in the top positions at the regional level. • A.4.1.5. Promote the monetization and benefits of data produced by the country's private sector, and its publication on the national open data platform. • A.4.1.6. Create a communication channel so that the academic sector, private sector, public or civil society in general, can request the publication of new open data, propose the publication of their data on the platform, and provide feedback on the quality of the published data

SO.4.2. To be a regional leader in the publication of data on biodiversity, native languages and other minorities in the country. • A.4.2.1. Create a bank of native languages, sign language and images to guarantee access for all citizens to public services (web, office, etc.) using translation and subtitling. • A.4.2.2. Create a database contextualized to each region of the country, for example: agro-industry, information on natural disasters, climate change, information on local

industries, socioeconomic information, etc. This database will be published on the national open data platform. • A.4.2.3. Promote the creation of a database of the biodiversity of our natural and cultural wealth, in addition to promoting the investigations of these databases. These databases will be published on the government's open data platform

E5 - Ethics

SO.5.1. To be a regional leader in the responsible use of data and AI algorithms

SO.5.1. To be a regional leader in the responsible use of data and AI algorithms. • A.5.1.1. Adapt the recommendations of the OECD "Principles on AI" (*), which Peru has signed, to the national reality, and start their implementation prioritizing research and development that stimulates the innovation of a reliable and accessible AI, prioritizing the sectors that the country considers key. •

A.5.1.2. Through the different regulatory bodies of public services and national superintendencies, evaluate an impact study on the less biased use of algorithms for the classification of people in the private sector. • A.5.1.3. Create an observatory to monitor, report on rankings indicators for the responsible use of AI such as the Oxford Insight Responsible AI ranking (**) and others. • A.5.1.4.

Create a unit to monitor and promote the responsible and ethical use of AI in the country.

SO.5.1. To be a regional leader in the responsible use of data and AI algorithms. • A.5.1.5. Implement a platform to be a registry of AI algorithms used in public sector. In addition to algorithms, the datasets used in use cases will be included. • A.5.1.6. In regulatory sandboxes created as a controlled environment where AI-based ventures can be promoted, promoting the ethical and responsible use of AI. • A.5.1.7. In the public sector, in all cases of use of AI to classify people (to provide benefits, opportunities or sanctions to citizens), they must have a socioeconomic impact study to guarantee equity. • A.5.1.8. Promote Ethics courses in all undergraduate and graduate programs in Computer Science, Software Engineering, and in all programs that contain AI courses

E6 - Collaboration SO.6.1. Facilitate the exploitation of synergies between universities and research centers through inter-university cooperation, the private sector, public organizations and international organizations SO.6.1. Facilitate the exploitation of synergies between universities and research centers through inter-university cooperation, the private sector, public organizations and international organizations. • A.6.1.1. Research projects presented to public funds such as

Fondecyt, Concytec and others, must include as evaluation criteria, the participation of at least two universities, one from Lima and the other from the provinces. • A.6.1.2. Promote collaboration with foreign universities, and include some criteria for the delivery of public funds to universities in CORE training programs, promoting the participation of prestigious universities in the US and Europe in master's and doctorates training programs. • A.6.1.3. Through a National Center for Innovation and Artificial Intelligence to promote collaboration in training, research, development and innovation of AI between the public and private sectors, academia, and alliances with prestigious foreign institutions. • A.6.1.4. Create and maintain a public registry of public, private and academic entities with AI/ML capabilities and conducting AI/ML research and development in the country. • A.6.1.5. Promote an alliance between countries in the region for the research and development of artificial intelligence. For example, explore a Pacific Alliance for Artificial Intelligence.

NATIONAL CENTER FOR INNOVATION AND ARTIFICIAL INTELLIGENCE

Purpose It is a national center of excellence in artificial intelligence, in response to the need for human talent and with the aim of accelerating the development and adoption of AI, which will carry out research, development, socialization and innovation by adopting Artificial Intelligence, coordinating its activities with the national and international academic sector, the private and public sectors, considering the sectors where Peru has competitive advantages, whether due to biodiversity, key economic sectors or criteria that are considered important for the country; inserting the country into a global value chain and promoting its integration into a globalized and digital world. It will prioritize the use cases where Artificial Intelligence generates solutions to the country's objectives such as the elimination of poverty, zero hunger, justice, quality education, health, clean and accessible energy, clean water and sustainable cities, the reduction of social gaps and others. use cases Functions • Carry out Research in Artificial Intelligence, as own projects or in collaboration with the academic, public and private sectors. • Carry out and promote the training of talent and skills at the national level in the research, development and adoption of AI framed within the ENIA. • Recommend and prioritize areas and cases of use of AI in coordination with the public, private and academic sectors of the country. • Carry out and facilitate AI job postings. • Sponsor national and international academic AI events.

NATIONAL HIGH PERFORMANCE COMPUTER CENTER

Purpose The National Center for High Performance Computing is the national center specialized in high performance computing (HPC) that will manage the country's super computers. This center will provide high-performance computing services to academia, the private sector, the public sector, and industry, and will also provide advanced training in high-performance computing, promoting their participation. The main players in this center will be universities and national research centers that will be connected to the centers through a national high-speed fiber optic network.

37. Guidelines for the use of artificial intelligence tools in the public sector

FROM: MINISTER SECRETARY GENERAL OF THE PRESIDENCY MINISTER OF SCIENCE, TECHNOLOGY, KNOWLEDGE AND INNOVATION TO: AS DISTRIBUTED

BACKGROUND

The swift advance of artificial intelligence has shown that in a very short period of time, it represents a pivotal transformer. In the face of this development, the Administration of the State must assume a proactive role in the management of this technology, understanding how it functions and utilizing it in the most beneficial way for society. This must always be done ethically, consciously, and responsibly, with an emphasis on people, and respecting the fundamental rights of those who work for the State. Empowering the State in the responsive use of artificial intelligence is fundamental, offering powerful reasons to enhance efficiency in public administration and to foster civic empowerment. Similarly, it is key to develop and strengthen national competitiveness at an international level, to encourage investment, innovation, and entrepreneurship in new technologies. This field, after all, involves issues that affect social, political, and economic dimensions. It is essential to build a common framework that encourages the use and development of technologies based on artificial intelligence within the different bodies of the Administration of the State, and at the same time, respect the rights of the persons in the face of the implementation of these types of tools. This framework will contribute, at the same time, to the goals of digital transformation of the State. As the Government of Chile has committed to the agenda of modernization and digital transformation for the public sector by the end of the year 2022, the Digital

Government Division (initially and subsequently, "DGD") of the Ministry of the Interior and Public Security (henceforth, the Ministry) will be in charge of coordinating the implementation of the Digital Transformation National Strategy, aligned with the regulation of the Law No. 21.180 of Digital Transformation of the State, coordinating the implementation of the Interoperability and Data Management System, implementing the State's Interoperability Network, and in conjunction with the Civil Protection Procedures and the Digital Readiness of the State, creating public policy 'ChileCompra', designing policies of Advanced Technologies and the Electronic Signature 'e-ChileCompra' and the 'Technological Compass for the Public Sector' ('TIC System').

5. In October 2021, the Ministry of Science, Technology, Knowledge and Innovation (hereinafter and interchangeably "MCTCI") published its National Intelligence Policy Artificial with the objective of inserting Chile into the global collaboration related to the development of tools based on this technology, having an ecosystem of research, development and innovation in the field that creates new capabilities in the productive, academic and state sectors, and that, following the principles transversal aspects of opportunity and responsibility, contribute to sustainable development and improve the quality of life for everyone. This instrument is currently in the process of updating.

6. In light of the above, it has been deemed necessary to recommend to the bodies of the State Administration the following guidelines regarding the use of tools of artificial intelligence in the public sector.

7. These guidelines will constitute a first step to arrive at a common framework, since provide general guidelines and recommendations to promote responsible, ethical, safe and transparent tools based on artificial intelligence systems in the public sector. Its monitoring will allow preliminary and detailed information about of the regulatory and coordination challenges at stake and, if necessary, determine What instruments are most efficient and effective to address them.

RECOMMENDED GUIDELINES

It is recommended to the bodies of the State Administration that, in the exercise of their functions, decide to develop, implement or use tools based on information systems. artificial intelligence, consider the following guidelines, through their respective

digital transformation coordinators:

Yo. Artificial Intelligence focused on people

to. Evaluate whether the use of artificial intelligence is the most appropriate technological solution both to meet citizen needs and good service and functioning of the Administration that they are trying to satisfy.

b. Ensure that people are not subject to arbitrary discrimination based on their race, color, ancestry, gender, sexual orientation, age, language, religion, political opinions, national, ethnic or social origin, economic or social condition of birth, disability or any other reason or reason, in accordance with the provisions by the Constitution or the laws, due to the use of tools of artificial intelligence.

c. Involve citizens and/or civil society, when appropriate, in the design and implementation of artificial intelligence tools, especially when they have the potential to affect their rights, through public consultations and other participation mechanisms provided for in national regulations.

d. Follow the guidelines of the Guide to the Ethical Formulation of Science Projects of Data, published by the DGD, for the implementation of these measures and others as well as the actions that derive from them, when applicable.

ii. Transparency and explainability

to. Inform citizens through suitable and easily accessible mechanisms about of the artificial intelligence tools that are being used within their processes, whether in the provision of public services or in supporting decision-making decisions of the State Administration body.

b. Communicate through suitable channels, easy to access and in clear language, if they use or they use artificial intelligence tools in their relationship with citizens, such as chatbots, virtual assistants or automated systems for customer service. public, among others.

c. Make available to citizens, through suitable channels that are easy to access and in clear language, information regarding the use of tools artificial intelligence to exercise its powers, especially in the case of decision-making processes that could affect fundamental rights of

people.

iii. Privacy and data use

to. Guarantee that the processing of personal data, especially those of a sensitive, which is made by using artificial intelligence tools, complies with the Law No. 19,628 on the protection of private life or its modifications, in particular, to ensure that data processed for the development, training or use of artificial intelligence tools are used exclusively for the purposes authorized by their owners or by law.

b. Do not enter personal information, especially that of a sensitive nature, in generative artificial intelligence tools, when these have not been contracted by the State Administration body or developed by or for this. In this regard, special care must also be taken with the information confidential information of legal entities to which the Administration has access.

c. Periodically review the studies, recommendations and jurisprudence of the Council for Transparency and other competent authorities in matters of protection of personal data, which are relevant for the implementation of these measures and the actions that derive from them.

d. Comply with the provisions of Decree No. 7, of 2023, of the General Secretariat Ministry of the Presidency, which establishes the Technical Standard for Information Security and Cybersecurity in accordance with Law No. 21,180; and Decree No. 11, of 2023, of the Ministry General Secretariat of the Presidency that establishes the Technical Standard of Quality and Operation of the electronic platforms that support procedures administrative bodies in the State Administration bodies and those relevant recommendations issued by the Ministry of the Interior and Security Public on cybersecurity.

iv. Other measures

to. Ensure, within the needs and possibilities of each service, training of its officials and other public servants of its organization for the adequate adoption of tools based on information systems artificial intelligence.

b. Pay attention to the guidelines, suggestions and learning contained in the following

documents:

Yo. Government study of data in the Ministry of Social Development and Family of the DGD and the Inter-American Development Bank.

ii. Guide "Allowed to innovate: How can we develop science projects of data to innovate in the public sector?" from the Government Laboratory.

iii. Publication "Transformative institutions 2023: good practices of the Central Administration" of the DGD.

8. For the purposes of following up on the recommendations made, the Head of the DGD, in the exercise of its functions, may directly officiate the bodies of the State Administration to request any data or information related to these guidelines. Likewise, the Director of Data of the DGD may request, via email, to the Digital Transformation Coordinators, any data or information related to this instrument.

9. The delivery of information from the Digital Transformation Coordinators in response to the data requested from the Director of Data of the DGD will be made at the email box Gobiernodatos~a digital.gob.cl.

10. When applicable, the Digital Transformation Coordinators will inform that another person from your institution has been designated to assume the duties of communication and follow-up described in this section. In that case, the Director of Data from the DGD must make all information requests to the new person that is designated.

11. The MCTCI, together with the DGD, within its scope of competence, may develop guidelines for the design and implementation of AI-based solutions in the public sector and facilitate compliance with the measures contained in this instrument.

12. Any change that may occur regarding the organizational structure of the Departments of the State mentioned in this instrument, will not affect in any way case the validity and validity of the expressed measures. Their roles defined in the present office will be assumed by their legal successors.

38. ChatGPT: How to use it in classes?

Strongly debated, the truth is that this tool is here to stay. Specialists agree that it simplifies and optimizes different tasks and jobs, both in the most diverse disciplines and in everyday life. How to use it in classes and take advantage of its advantages? Here are some recommendations.

What are you? A useful tool or a threat? We asked ChatGPT himself: "I'm an artificial intelligence guy. "My main purpose is to process and generate text in response to user questions and requests," he tells us and adds that yes, it is useful," he replies.

"I am designed to be a useful tool. My main purpose is to help and provide information on a wide range of topics. I can assist in text generation, answer questions, offer suggestions and provide support on various tasks," she adds.

However, ChatGPT emphasizes that "my usefulness depends on the context and the specific needs of each user. It is important to remember that I am a tool and that it is always advisable to use human judgment and consult other sources when necessary to make informed decisions.

ChatGPT – if we use version 3 – "in practice, and in more general terms, is an artificial intelligence chatbot (virtual operator) created by the artificial intelligence research laboratory Open AI, and based on a large language model "specially trained to establish conversations," says Gabriela Arriagada, researcher at the Institute of Applied Ethics and the Institute of Mathematical and Computational Engineering.

As specified in the document "A brief look at the current state of Artificial Intelligence" , prepared by professors Marcelo Arenas , Gabriela Arriagada , Marcelo Mendoza and Claudia Prieto , the GPT Chat "was trained to answer questions and generate answers in natural language . One capacity that Chat GPT has is to incorporate a conversation into the text generative mechanism, placing emphasis on the feedback obtained in the interaction with the human." Users can ask you questions and ask you to prepare organized information to solve requests, such as a math problem or an essay about a book.

A threat?

So, is ChatGPT a threat? Professor Gabriela Arriagada explains that the way it works today, in versions 3 and 4, does not present a specific threat. As a tool, it simplifies and optimizes a series of tasks that can facilitate the work of different disciplines.

"It seems to me that the arrival of a tool that helps us optimize information searches, that helps us structure lists, make summary documents, review codes, etc., is welcome, as long as we understand what it does, how "It does it, and how to translate its use into a complement to our work and not a supplement," says the professor.

However, it must be considered, as the aforementioned document warns: "A limitation that Chat GPT has is that it requires retraining to incorporate updated information. This limits their ability to interact around current issues." Although version 4, launched last March, has "a greater knowledge base to interact around current issues. It is also capable of handling longer contexts (...) favoring the creation of content, extensive conversations and automatic document analysis." As Engineering professor Marcelo Mendoza warns, on the one hand "it is an opportunity, because it produces an enormous advance in language technologies and represents a significant advance in the state of the art of natural language processing." But at the same time it is also a threat, "in the sense that there are many aspects of GPT Chat that have not been studied, such as, for example, its ability to create biased content or produce disinformation."

How to use ChatGPT in classes?

Surprise was what the Institute of Astrophysics professor Thomas Puzia saw on the faces of his students when he asked them to use tools like ChatGPT for the Astrobiology course.

"I wanted to motivate them, for this tool to be a support to search for references, review titles of books and articles, to make summaries of a lot of literature. This was a tremendous advantage for them. It was noted that the understanding of the material improved. I also promoted the use of artificial intelligence to prepare slides for an exhibition, which also improved the quality of their presentations," says the professor.

The academic took into account that the use of this technology did not have a "good reputation", but he proposed that this was not copy and paste, but rather he explained to the students that this was to understand, verify, practice logic and self-criticism of knowledge. In fact, he says his performance improved and his world expanded.

"With these tools, students can dedicate much more time to things that are more important, such as understanding concepts, their causality with each other, among other things. Without a doubt it has been a tremendous incentive for learning. This is the new world and it has arrived," he says.

What does the UC propose?

As Chantal Jouannet , director of the Teaching Development Center, explains, ChatGPT can have very interesting uses to assist us in the areas of teaching management and generation of learning experiences. On the one hand, she says, it can be used as an assistant to facilitate teaching tasks, such as designing questions for the class, defining performance indicators, designing evaluation instruments, and more.

This, always using teaching criteria to analyze the relevance of the chat responses. "On the other hand, it can be used to design specific active learning strategies and in evaluations. For example, we can ask students to analyze and criticize a text prepared by ChatGPT 3. The chat can be used as a counterpart in a debate and can even be used as an object of study when discussing its operation and the ethical implications of its use. , Explain.

As Professor Gabriela Arriagada adds: "The basic criterion is to be clear about the purpose of its use, what am I going to use it for? Using it as an assistant is usually a good motivation, so being transparent in its use is important. "A good practice is always to declare what has been used and for what." And she specifies: "Normalizing the declaration of the use of generative tools is the best way to combat the risks of plagiarism, originality, or distrust and fear of the integration of technology as a support tool."

In the midst of the debate over the integration of artificial intelligence in the educational world, last May the Ministry of Education launched the new "Teaching guide: how to use ChatGPT to enhance active learning" . This document was born as a way to anticipate and address the challenges of new technologies to take advantage of the opportunities that artificial intelligence applications, such as ChatGPT, present in education.

"Another relevant aspect is maintaining a critical stance towards technology , " says Gabriela Arriagada. And she adds: "This means understanding that, although Chat GPT is a great advance in contrast to previous technologies, we cannot trust that everything it shows us is perfectly acceptable."

Chantal Jouannet highlights that the most important thing is to recognize that these tools have great positive potential in university education, just like other technologies in other times. However, she emphasizes the importance of its responsible use: “It is essential that there be dialogue with the student body about the contributions and limits of its use, as well as making it transparent when the performance of an activity has been carried out with the support of ChatGPT,” she explains. To investigate this topic further, both for students and academics, the Center for Teaching Development, the Academic Directorate of Teaching and UC Libraries, jointly developed a site with plenty of material and guidelines for the use of artificial intelligence tools, such as ChatGPT at UC , which can be reached here.

What is the future of ChatGPT?

“This technology is already embedded in services that we use every day. It is already connected to knowledge bases, so it will serve as an assistant for news writing, veracity verification (automatic fact-checking) and programming assistant along with tools such as copilot,” comments Professor Marcelo Mendoza.

As he adds, “the next step in these technologies is the generation of multimodal content . Currently the generation is unimodal, that is, you either generate text (GPT Chat) or you generate images (Dall-E). The production of bimodal content - for example, text + images - will allow the production of content such as concept maps or other content that combines text and images.

“It is possible that the quality of the responses, interactions, and results that these models and their chatbot counterparts can give us will be much more refined. It would not be unreasonable to see better chatbots implemented for customer services or to help with organizational tasks,” says Gabriela Arriagada.

“A future with supporting technology like this can change the way we work – especially if this is adopted massively – freeing us from simpler tasks, but demanding more of us with respect to more elaborate tasks ,” says the researcher and concludes: “We should not, however, talk about a future only of advances, but also of regulations and standards that support these innovations.”

39. International AI Cooperation and Governance Forum 2022

International AI Cooperation and Governance Forum 2022

Theme: AI Governance and International Cooperation for a Resilient Future

Date: December 9th - 10th, 2022

Form: Online

Host: Tsinghua University

Organizer: Institute for AI International Governance of Tsinghua University (I-AIIG)

Supporting Partners:

United Nations Development Programme

United Nations Educational, Scientific and Cultural Organization

UN Women

International Labour Organization

Institute for Ethics in Artificial Intelligence, Technical University of Munich

Center for Humanitarian Dialogue;

International Research Center for AI Ethics and Governance, Institute of Automation, Chinese Academy of Sciences

The Chinese Institute of New Generation Artificial Intelligence Development Strategies

China Academy of Industrial Internet

Institute of Artificial Intelligence, Tsinghua University

Institute for AI Industry Research, Tsinghua University

Center for International Security and Strategy, Tsinghua University

Center for AI Governance (Beijing)

China Institute for Science and Technology Policy, Tsinghua University

Center for Science & Technology Development and Governance, Tsinghua University

Tsinghua University Latin America Center

Tsinghua Southeast Asia Center

SPARK UNDP SDG Innovation Lab | Chengdu

National Communication Center for Science and Technology, CAST

Background and Objective

In recent years, artificial intelligence (AI) represented by super large-scale pre-training model grows by leaps and bounds with more resources of data and computing power, injecting intelligent impetus to research in life science, mathematics, space science among other fields. Meanwhile, expanded application scenarios also nourish the development of AI, driving traditional industries to go intelligent. In particular, against the urgent needs for global economic recovery in the post-pandemic era, AI plays a more prominent role in empowering scientific research and innovation and industrial development, and underpins improvement of capacity of risk prediction, crisis response and recovery, adding to the resilience of the social and economic development.

However, as AI technologies are more deeply embedded in economic and social life, transforming our means of production and lifestyle, it also elicits more prominent challenges in security, privacy and equity. Over the past year, governments, international organizations and industrial communities have rolled out a series of norms and regulations, policies and standards of AI governance, making AI governance one of the most important agendas in global governance.

At the just-concluded 17th G20 Leaders' Summit, China's President Xi Jinping proposed an initiative aimed at making global development more resilient. The Global Development Initiative that China proposed is an innovation-driven approach. The historic opportunity of recent revolutions in science and technology, and industrial changes should be seized to accelerate the transformation of these technological achievements to productivity. It is important to foster an open, fair, just, and non-discriminatory environment for science and technology development. At the same time, cultivate new momentum for economic growth to pursue leapfrog development together.

China has recently formally submitted its Position Paper on Strengthening Ethical Governance of Artificial Intelligence (AI) to a meeting of the States Parties to the UN Convention on Certain Conventional Weapons. Based on the implementation of the policies on technology ethics in China and useful international experience, the Paper highlighted the priority of ethics in AI governance, and proposed systematic recommendations for strengthening the ethical governance of AI, with emphasis on promoting the establishment of an international framework and standards for AI governance based on broad consensus through international exchange and collaboration.

In this context, the International AI Cooperation and Governance Forum 2022 will be convened under the theme of "AI Governance and International Cooperation for a Resilient Future". Focusing on regulation, R&D, utilization and international cooperation of AI, it explores the establishment of a governance system appropriate for the sound development of AI; development of governance technologies and tools to guarantee responsible use of AI; and the best practices in AI governance of the political, industrial, academic and research communities in different countries and regions. The forum aims to promote AI to provide new momentum for and unleash potential of scientific innovation and industrial development, so as to make global growth more inclusive, resilient and sustainable for the common well-being of humanity, and for a community with a shared future for mankind.

40. Responsible Use of AI – Guidance from a Singapore Regulatory Perspective

Responsible Use of AI – Guidance from a Singapore Regulatory Perspective

I. Introduction

Artificial intelligence ("AI") is no longer a mere concept of the future. Recent developments in AI technology have opened the doors to a wide range of practical use cases. This has been swiftly adopted by the commercial world across a variety of business functions, with the accelerating uptake rate indicating that AI systems are set to become ever more prevalent in our daily lives. As with any newly adopted technology, AI brings with it certain issues and concerns, which are further exacerbated by a general lack of familiarity. These risks have been brought to the forefront in light of the recent popularity of AI solutions, including issues of ethics, mistakes and hallucinations, privacy and confidentiality, disinformation and cyber-threats, and intellectual property.

A common theme across AI adoption is the responsible use of AI – how should AI solutions be implemented, what forms of testing are available for AI systems, and what are the best practices when using AI? In the absence of established standards and practices, businesses have been looking to industry regulators for guidance.

In this regard, Singapore regulators have demonstrated their awareness of and proficiency with AI and its related risks. In recent months, Singapore regulators have provided guidance on the responsible use of AI for businesses in various industries. These assist businesses utilising AI tools or seeking to implement such tools, and provide an indication of how AI regulations may be structured when established.

In this article, we take a look at some of these initiatives in the Singapore context:

The launch of the AI Verify Foundation, which aims to develop the AI Verify testing tool for the responsible use of AI;

The public consultation on the Proposed Advisory Guidelines on Use of Personal Data in AI Recommendation and Decision Systems; and

The Veritas Toolkit version 2.0 for the responsible use of AI in the financial sector developed by the Monetary Authority of Singapore ("MAS").

II. AI Verify

On 7 June 2023, the Infocomm Media Development Authority ("IMDA") announced the launch of the AI Verify Foundation ("Foundation"), which has the aim of harnessing the collective contributions of the global open-source community to develop the AI Verify testing tool for the responsible use of AI. The Foundation will look to boost AI testing capabilities and assurance to meet the needs of companies and regulators globally.

The Foundation will:

Foster a community to contribute to the use and development of AI testing frameworks, code base, standards, and best practices;

Create a neutral platform for open collaboration and idea-sharing on testing and governing AI; and Nurture a network of advocates for AI and drive broad adoption of AI testing through education and outreach.

AI Verify provides organisations with an AI Governance Testing Framework and Toolkit to help validate the performance of their AI systems. Furthermore, AI Verify is extensible so that additional toolkits, such as sector-specific governance frameworks, can be built on top of it.

AI Verify is a single integrated software toolkit that operates within the user organisation's enterprise environment, facilitating the conduct of technical tests on the user's AI models and the recording of process checks. AI Verify's testing processes comprise technical tests on three principles: fairness, explainability, and robustness. Process checks are applied to the identified principles. In recognition of global compliance requirements, the testing framework is consistent with internationally recognised AI governance principles, such as those from the EU, OECD and Singapore.

The development of AI Verify and the launch of the Foundation indicate the Government's recognition of the importance of tools that are able to adequately test the performance of AI systems. Organisations using AI in their businesses require more reliable and standardised test systems, which will then allow them to make provisions for protection from the resulting risks.

III. Advisory Guidelines

The Personal Data Protection Commission ("PDPC") has launched a public consultation ("Consultation") seeking views on the Proposed Advisory Guidelines on Use of Personal Data in AI Recommendation and Decision Systems ("Guidelines"). The Consultation has ended on 31 August 2023.

The aim of these Guidelines is to:

Clarify how the Personal Data Protection Act ("PDPA") applies to the collection and use of personal data by organisations to develop and deploy systems that embed machine learning models ("AI Systems") which are used to make decisions autonomously or to assist a human decision-maker through recommendations and predictions; and

Provide baseline guidance and best practices for organisations on how to be transparent about whether and how their AI Systems use personal data to make recommendations, predictions, or decisions.

The use of personal data in AI Systems raises important issues of privacy and confidentiality. Personal data may be used in the training of various AI Systems, including AI recommendation and decision systems in e-commerce to recommend and personalise products or content to users, and AI tools to predict product demand. While such data may be essential to the training process, organisations must be aware of how it interacts with their data protection obligations under the PDPA. The breach of such obligations may lead to the imposition of potentially onerous penalties and fines, as well as reputational damage. The Guidelines will thus be a vital source of guidance in this regard.

IV. Responsible use of AI in the financial sector

On 26 June 2023, MAS announced the release of the Veritas Toolkit version 2.0, an open-source toolkit to enable the responsible use of AI in the financial industry, by helping financial institutions ("FIs") carry out the assessment methodologies for the Fairness, Ethics, Accountability and Transparency ("FEAT") principles. This is part of the MAS' Veritas Initiative which was first

announced in November 2019. The Veritas Toolkit version 2.0 builds on the earlier Veritas Toolkit version 1.0 which had been released in February 2022 that focused on the assessment methodology for fairness. The Veritas Toolkit version 2.0 features an improved fairness assessment methodology and new assessment methodologies for ethics, accountability and transparency. The FEAT principles provide guidance to firms offering financial products and services on the responsible use of AI and data analytics. The Veritas Toolkit is the first responsible AI toolkit developed specifically for the financial industry.

In addition, the consortium behind the development of the Veritas Toolkit has published a white paper setting out the key lessons learnt by seven FIs which piloted the integration of Veritas methodology with their internal governance framework, including the importance of:

- a consistent and robust responsible AI framework that spans geographies;
- a risk-based approach to determine the governance required for the AI use cases;
- responsible AI practices and training for the new generation of AI professionals in the financial sector.

The MAS also announced additional use cases which the consortium had developed to demonstrate how the toolkit could be applied, including the application of transparency assessment methodology to predictive AI-based policy underwriting for insurers as well as application of the FEAT assessment methodology to fraudulent payment detection systems.

MAS has stated that the consortium will focus on training in the area of responsible AI and facilitate the adoption of the Veritas Methodologies and Toolkit by more FIs.

In line with MAS' focus on responsible use of AI, on 31 May 2023, MAS and Google Cloud signed a Memorandum of Understanding ("MoU") to collaborate on generative AI solutions grounded on responsible AI practices. The MoU provides a framework for cooperation in technology and industry best practices in three areas:

- Identifying potential use cases, conducting technical pilots, and co-creating solutions in responsible generative AI for MAS' internal and industry-facing digital services;

- Cooperating on responsible generative AI technology application development and test-bedding of AI products for business functions and operations; and

- Supporting the technical competency development on responsible generative AI and deep AI skillsets for MAS technologists.

V. Concluding Remarks

Singapore regulators such as IMDA, PDPC and MAS have demonstrated themselves to be deeply involved in issues of AI deployment and development and how they apply to their respective industries. This can be seen by their efforts at developing toolkits and guidance papers for organisations and businesses on the responsible use of AI.

Currently, it remains to be seen whether specific AI legislation or regulations will be developed to impose binding obligations on AI users. In the meantime, the guidance offered by the regulators in the initiatives highlighted above may provide a shape of things to come.

41. Regarding the Use of Generative AI

Regarding the Use of Generative AI

To All Students,

Generative AI such as ChatGPT has been developing rapidly in recent years, and the use of generative AI, such as those that generate language, and even images and voice, is expanding throughout all of society. Nagoya University's views on generative AI are described in "[The Use of Generative Artificial Intelligence Technology \(Generative AI\) in Education and Research](#)". However, we ask all students to pay special attention to the following.

Please think about the effects and negative aspects that the use of such AI may have on your own learning, and deal with generative AI with the sincere intention of deepening your own learning. Please follow the guidance of your instructors regarding the use of generative AI in classes.

Please remember the following points.

1. There is a possibility that information provided by generative AI may contain errors, and the use of such misinformation may result in disadvantages for the user. Therefore, it is necessary for you to examine the accuracy of the provided information yourself.

2. Even if the provided information is correct, it may cause copyright infringement or plagiarism issues. Therefore, information provided by AI should only be used as a reference for your writing. Please do not use the provided information as is, even only partially. Penalties may be imposed in such cases.
3. Information entered as questions when using generative AI has a risk of being leaked, so you must not enter information that should be kept confidential, such as personal information.

42. Guidelines for Instructors Regarding AI in University Education

Background

ChatGPT, an AI (artificial intelligence) text generation system developed by OpenAI, was released in November 2022. It can answer questions in fluent natural language based on a vast amount of learning data, and can also create summaries and outlines in addition to generating text. It is widely accessible on the web and can be used in a conversational format, leading to explosive growth in its usage and having a significant impact on university education. This “Guideline” aims to provide a broad guideline for the use of AI in educational settings at this university, taking into account machine translation tools such as DeepL as well as text generation AI such as Bing Chat, with a focus on ChatGPT and similar text generation AI. Please note that the contents of this “Guideline” reflect the current situation, taking into consideration that ChatGPT is still under development and various services that incorporate ChatGPT or other conversational text generation AI (such as Bing Chat) have been released. It is also important to note that this “Guideline” is not intended to provide guidance for academic evaluation purposes.

How AI Generates Text AI text generators work based on large language models (LLM)

LLMs are trained on a massive amount of natural language text using deep learning, which predicts the most likely words to follow a given sequence of words as input. This input is called a prompt, and if a partial sentence is input as a prompt, the AI can complete the sentence by predicting the most likely words to follow. It is also possible to input a question asking for an explanation of something, and the AI can output a highly probable sentence that answers the question. However, even with the same question, the output answer may differ due to subtle differences in the prompt, such as the position of a comma. It is also known that the accuracy of the answer can vary depending on the prompt, and prompt engineering, which is a method for obtaining highly accurate answers, is attracting attention. However, the text generated by

AI is based on the results of machine learning from given language data, so it is not necessarily accurate all the time. It has also been pointed out that undesirable expressions may be included in the output. Therefore, in new AI systems, machine learning (reinforcement learning based on human feedback) has been introduced to encourage the selection of more desirable responses. Based on human evaluation according to the 3H criteria (helpful, honest, harmless), the multiple candidate responses that the system is attempting to output are ranked in the order of desirability. With these improvements, the overall accuracy of text generation AI is improving every day, but since it is not necessarily possible to obtain complete answers or avoid mistakes, users need to be mindful of appropriate use.

Limitations of AI Text Generation

There are several limitations to AI text generation that are known:

1) AI may provide incorrect answers, even if they are different from the facts. It is always necessary to check the accuracy of the information by returning to reliable sources when considering AI responses. Additionally, AI may have difficulty in certain fields, such as providing inaccurate answers to mathematical problems. 2) AI responses may contain potential biases or errors. Since AI is trained based on data found on the internet, there is a possibility that biases and errors found on the internet may be inherited by AI responses. Additionally, the unconscious biases of evaluators may also affect AI responses during the reinforcement learning process. 3) In the case of ChatGPT, its knowledge is based on data learned from the internet up until September 2021, and it cannot answer questions about events that have occurred since then. However, Bing Chat is connected to the internet and therefore, can provide answers about the latest events. 4) ChatGPT cannot display the sources used for its responses, whereas Bing Chat can display the internet sources used for its responses.

Guidelines for the use of AI in university education

AI is undoubtedly a powerful tool for assisting students in writing, despite its limitations mentioned above. ChatGPT (or various other AIs that may emerge in the future) seems to have become a tool that students will be expected to master when they enter society. Therefore, in universities, it is necessary to educate students on appropriate usage and ways of thinking in the context of learning, rather than limiting their use. It is important to provide a balance between the

philosophy and practice of educating students on the appropriate and effective use of AI while ensuring that AI use does not deviate from the primary purpose of university education, and that it helps to enhance students' critical thinking skills. As for the use of AI in individual classes, the instructor should be allowed to determine whether to prohibit, restrict or actively utilize it, depending on the characteristics of the course. However, regardless of how it is used, the following points should be considered when implementing AI in the classroom:

- Understanding of AI: Instructors must have a basic understanding of AI. If they have not yet tried using AI, they should try it immediately. Especially when it comes to assignments that are planned to be presented in class, it is desirable to confirm in advance what kind of answer AI will provide.
- Shared understanding with students: At the beginning of the semester, instructors should explain to students their own thoughts on the use of AI in the class and engage in ongoing discussions with them on how AI can be involved as a rule in achieving the purpose of the course and its process.
- Clear rules: At the same time, instructors should establish clear rules for the use of AI in the course for their students. In particular, when presenting assignments during the semester, it is recommended to reconfirm the rules with the students.
- Fairness of rules: When establishing rules for the use of AI in class, please be aware of ensuring fairness among students. Some students may not be good at using computers while for other students it may be indispensable to use them (due to disabilities, for example). Please give sufficient consideration to the characteristics of each student.
- Ethical use of AI detection services: There are services that can detect whether an article was generated by AI or not, but please be aware that there are cases of false positives (misidentifying articles written by humans as AI-generated articles) and false negatives (misidentifying AI-generated articles as human-written articles). Therefore, it is not recommended to fully rely on the results of these services.
- Use of AI and personal information: When using AI for questions and assignments, instructors should take measures to ensure that students' personal information is protected. It is important not to ask questions that include personal information because information included in questions addressed to AI will be incorporated in the AI database and may be used later to answer questions by a third party

Resources for Understanding and Using Article Generation AI

The website of OpenAI that developed ChatGPT. You can register here to use ChatGPT.
<https://openai.com/> Bing Chat, a search engine incorporating ChatGPT provided by Microsoft. To use this

service, you need to register using Microsoft's browser, Edge. <https://www.bing.com/> “ChatGPT: AI Education-related Information Summary” by Professor Rui Yoshida of the University of Tokyo, which summarizes the latest education-related information in Japanese. <https://edulab.t.u-tokyo.ac.jp/chatgpt-ai-resources/> In particular, Professor Yoshida’s video “Let’s talk about using ChatGPT in education” is a useful reference for educators considering using ChatGPT in the classroom. <https://edulab.t.u-tokyo.ac.jp/2023-02-11-report-event-chatgpt/> “The Evolution of AI and Japan’s Strategy” by Professor Yutaka Matsuo of the University of Tokyo, provides a concise summary of the history of AI development leading up to ChatGPT and how AI works.

<https://note.com/api/v2/attachments/download/a29a2e6b5b35b75baf42a8025d68c175>

43. Use of Artificial Intelligence Tools in Teaching, Learning and Assessments

Artificial intelligence (AI) is sweeping the globe, and generative AI in particular has been hotly discussed because of its potential to revolutionize the way we teach and learn. The University believes that it is crucial for teachers and students to embrace and become acquainted with AI in order to optimize its potential in education. Students should learn to make sensible use of AI tools, not only for their studies, but also for their future professional development and advancement in order to thrive in this AI era. Like any other educational resources, teachers and students should approach AI tools critically, recognising their limitations in an honest and authentic manner and how these tools could be incorporated into teaching and learning in order to attain the desired learning outcomes. To optimize the use of AI in education, the University has prepared for teachers the “Guidelines on the Use of Artificial Intelligence Tools in Teaching, Learning and Assessments” (Guidelines) to i) set out how the University may integrate AI tools in its teaching and learning while upholding academic honesty, integrity and quality; ii) recommend some possible approaches in adopting AI tools in teaching and learning; iii) make it clear and explicit that improper/unauthorized use of AI tools in assignments/assessments constitute acts of academic dishonesty which will be handled in accordance with the University’s existing guidelines and procedures; and iv) provide some readily available references and resources for supporting the adoption of AI tools in teaching and learning.

Students should take note of the following salient points extracted from the Guidelines and follow strictly the instruction and/or permission given in the course outline by the teachers regarding the

use of AI tools in teaching, learning and assessments. The Guidelines will be reviewed and updated as needed to reflect changes in technology, best practices, and other relevant developments.

- a) There are different types of AI tools, for instance generative AI tools (e.g. Chat GPT) which can be easily instructed using ordinary human language to generate various formats of texts. Some AI tools facilitate the creation of ‘original’ artwork (e.g. DALL·E 2), translated text (e.g. Google Translate), formulas (e.g. Sheet+), and computer code (e.g. OpenAI Codex), etc. applicable to a great variety of use. While teachers and students are encouraged to explore and take advantage of the benefits of adopting appropriate AI tools to enhance their teaching and learning activities, decisions on which AI tools to adopt and how to use them in teaching and learning should be made cautiously and thoroughly.
- b) The availability and accessibility of AI tools to students will be carefully evaluated before adopting AI tools into any teaching and learning activities. AI is a double-edged sword; we should use but not abuse it, use it as a research but not cheating tool, and most importantly, use AI to think with you, but not for you.
- c) Where applicable and permitted, approaches to the use of AI tools in different disciplines will be worked out taking into consideration the needs of different disciplines, their pedagogical approaches and assessment means. When adopting AI tools in teaching and learning, teachers and students should be cautious of their accuracy and reliability and bear the responsibility of using the educational resources and references obtained through these tools.
- d) As a general principle, students are prohibited from using any AI tools to complete their assignments, assessments and any other works that count towards their final grade of the course or attainment of the desired learning outcomes, unless explicitly permitted
- e) Depending on the learning outcomes, pedagogical design, and assessment scheme of different courses, the following are some possible approaches to adopt AI tools in teaching and learning. Relevant details will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction and permission strictly and seek clarification from the course teacher if in doubt. Students are also expected to understand the limits and appropriate uses of these tools.
 - i. Approach 1 (by default) - Prohibit all use of AI tools In assessing the level of achievement of learning outcomes and students’ performance, students are expected to produce their own work independently without any collaboration or the use of AI tools. That says students are prohibited

from using any AI tools in their assignments and assessments that count towards students' final grade of the course, or for evaluating their attainment of the desired learning outcomes.

ii. Approach 2 - Use only with prior permission In some courses, it may be appropriate to use AI tools in some in-class exercises or assignments. Where applicable and permitted, students will be clearly and explicitly informed of when and how they can use these tools which shall be cited or acknowledged in their work. Details will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction and permission strictly and are expected to understand the limits and appropriate uses of these tools.

iii. Approach 3 - Use only with explicit acknowledgement In courses where students are allowed or expected to collaborate with or use AI tools, students may use these tools for in-class learning activities, exercises or assignments as long as they explicitly cite or acknowledge the use of these tools. Details will be spelt out clearly in the course outline and/or the instructions of the assignments. Students shall follow the instruction strictly and are expected to understand the limits and appropriate uses of these tools.

iv. Approach 4 - Use is freely permitted with no acknowledgement In courses where students are allowed or expected to frequently collaborate with or use AI tools, students may use these tools for in-class learning activities, exercises or assignments without citing or acknowledging the use of these tools. In these classes, it is critical that students understand the limits and appropriate uses of these tools. Details on which AI tools are to be used will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction strictly and are expected to understand the limits and appropriate uses of these tools

f) The adoption of permitted use of AI tools in courses are subject to regular review by the course teacher(s) and the programme concerned. It is the responsibility of students to study the course outline, assessment scheme and instruction of individual assignments in detail to ensure that they follow the instruction and permission strictly. Improper/unauthorized use of AI tools in learning 3 activities and assessments constitute acts of academic dishonesty which will be handled in accordance with the University's Procedures for Handling Cases of Academic Dishonesty.

g) As a general principle, students are expected to complete assignments/assessments on their own without any external assistance, unless otherwise specified. If AI tools are permitted for use in the course, students should pay attention to the following for proper use of these tools:

- i. students should learn and use these tools responsibly and ethically, and be aware of their limitations;
- ii. the quality of output of some generative AI tools correlates directly to the quality of input, students

should master “prompt engineering” by refining their prompts in order to get good outcomes; iii. students should fact-check all outputs of AI tools by cross-checking the claims with reliable sources and are responsible for any errors or omissions, if any, when using these tools; iv. like any other tools and references, permitted use of AI tools should be acknowledged unless otherwise specified; specific and detailed information on the AI tools used, including prompts used if applicable, for completing the assignments/assessments should be provided in the work concerned and, if deemed necessary, the output of generative AI should be included as an appendix of the work submitted by students.

h) As a general principle, students are expected to complete all coursework, formative and summative assessments independently without the use of AI tools or other forms of unauthorized assistance, unless specifically permitted. Improper and unauthorized use of AI tools not only jeopardize the quality and efficacy of teaching and learning, but they also constitute acts of academic dishonesty. Students should be cautious of the following which may result in improper/unauthorized use of AI tools in learning: i. using AI tools in completing assignments/assessments without prior permission; ii. handing in an AI-generated work as one’s own; iii. using AI tools to cheat in a course; iv. using AI tools that are not up to date and result in the use of outdated and inaccurate resources; and v. using AI tools in an unethical and irresponsible manner.

i) If the use of AI tools is not permitted in an assignment/assessment, and a student is later found to have used such a tool in the assignment/assessment, the case should be handled in accordance with the University’s Procedures for Handling Cases of Academic Dishonesty

j) Similar to the submission of other assignments, students are required to declare and assure that the works submitted are their original works except for source material explicitly acknowledged, and the permitted use of AI tools in the assignment(s), if applicable. The academic honesty declaration statement is updated accordingly in the VeriGuide System.

k) While appropriate enhancement would be implemented to the VeriGuide system for detecting cases of improper/unauthorized use of AI tools, the University will explore and devise other appropriate measures in detecting students’ submissions and assessments. All identified cases will be handled in accordance with the University’s Procedures for Handling Cases of Academic Dishonesty.

- l) Similar to other serious cases of academic dishonesty, penalties for improper/unauthorized use of AI tools in assignments/assessments may include reviewable/permanent demerit(s), failure grade for the course concerned, suspension from the University, lowering the degree classification, and termination of studies at the University.
- m) For cases in which improper/unauthorized use of AI tools in assignments/assessments was committed by a person when s/he was still a student of the University studying for the award in question (but only came to light after graduation) may also constitute good cause under which the University may revoke his/her academic award in accordance with the University's procedure for revoking an academic award
- n) The educational needs of students in the use of AI tools will be reviewed.
- o) University will provide professional development opportunities and support to teachers to enhance their AI literacy so as to help enhance sensible and critical use of AI tools among students.
- p) Guidelines and policies on the use and application of AI tools in teaching and learning will be regularly reviewed and updated to reflect changes in technology, best practices, and other relevant developments.

44. Seoul National University AI Policy Initiative

Introduction Past Activities Seoul AI Policy Conference 2017–2020 2017 Conference 2018 Conference 2019 Conference 2020 Conference (Virtual) 2020 COVID-19 Webinar SAPI Book Publishing Event – Data Ownership: Who Owns My Data? SAPI Issue Paper Seminars AI and Future Society Media Algorithm and Democracy Data Privacy Law Amendments: Pseudonymization, Research Utilization, Data Transaction 04 08 16 18 20 Other Academic Events Book Publication DAIG (Data & AI Governance) Magazine Publication Issue Paper Publication 24 30 28 32 Table of Contents Current Research Activities: Projects Other Activities 34 38
Introduction: Various policy issues are being raised due to the rapid development of artificial intelligence technology and social changes resulting from its adoption. SNU AI Policy Initiative (SAPI) reviews and analyzes social challenges that the data-driven artificial intelligence technology will bring. SAPI aims to address these issues through interdisciplinary research by combining disciplines such as technology, humanities, social science, and law. SAPI launched in April 2017 with the vision of becoming a ‘Social Lab’, a research platform for various disciplines within the University. SAPI is leading research in the field of artificial intelligence & law and is

actively engaged in joint research and cooperation with domestic and international experts in fields related to data and artificial intelligence.

Seoul AI Policy Conference 2017–2020: At the 1st Seoul AI Policy Conference in 2017, the panel discussed the algorithm technology and privacy policy. The opening session introduced major policy issues regarding artificial intelligence, focusing on algorithm technology and privacy issues. Session 1 panels discussed policy issues of artificial intelligence and big data in the context of competition, and Session 2 panels discussed policy issues pertaining to the algorithm technology. Session 3 panels focused on the policy issues regarding privacy and data protection. In 2018, the panels discussed the theme of accountability and regulation in the era of artificial intelligence. The first session mainly focused on the regulatory issues and discussed the issue of establishing proper data governance. The second session focused on accountability and discussed the growing ethical concerns regarding the proliferation of automated decision-making. The third session focused on other technical issues such as blockchain and discussed the proper governance of the new technologies. In 2019, the conference theme was ‘the implementation of trust and fairness in the age of artificial intelligence’. In the first session, the panels discussed the reliability of artificial intelligence and the governance framework necessary to gain public trust. In the second session, the panels discussed what it means for the machine to be “fair” and discussed the approaches taken to implement fairness in artificial intelligence. Unlike in the previous years, the 2020 conference was held virtually due to the COVID-19 pandemic. Accordingly, SAPI expanded the number of sessions to eight and arranged the sessions on different days so that sufficient discussions can take place. Despite the change of format, the invited presenters of all sessions gave insightful presentations. For instance, Professor Cynthia Dwork delivered the keynote lecture on current state of Differential Privacy. Despite being an online event that spanned over several weeks, hundreds of scholars attended every session.

2020 COVID-19 Webinar: Data, AI Governance, and COVID-19: Medium and Long-Term Perspectives for Asia SAPI, jointly with Singapore Management University and University of Tokyo held the conference based on the “Data, AI Governance, and COVID-19: Medium and Long-Term Perspectives for Asia” report published in September 2020. Researchers from Korea, China, Japan, and Singapore discussed the initial COVID-19 response of these Asian countries in terms of data and AI governance. Researchers further discussed how COVID-19 responses will

transform our societies. The research was conducted as a part of SMU-Microsoft Asian Dialogue on AI Governance. (<https://caidg.smu.edu.sg/smu-microsoft-asiandialogue-ai-governance/events>) SAPI Book Publishing Event: Data Ownership: Who Owns My Data? To commemorate the publication of ‘Data Ownership’, SAPI organized a book publishing event on April 29, 2019 with the authors. The data economy and artificial intelligence technology all stem from data. The question that is constantly being asked is, ‘Who owns this data?’ The event was a place for discussion to organize various thoughts on this question and seek policy answers. AI and Future Society At this event held on May 16, 2019, panels discussed how artificial intelligence affects economic growth, what implications the transition to the data economy has on market competition, what changes are happening in the labor market. Panels further evaluated the efforts being made in Korea to establish AI principles that deal with issues such as fairness and discrimination.

SAPI Issue Paper Seminars Media Algorithm and Democracy At the event, held on November 17th, 2019, panels discussed how algorithms or user profiling work and how they are affecting democratic decision-making process. Panels evaluated phenomena often referred to as the echo chamber effect or filter bubble and how to approach algorithmic transparency and explainability with respect to algorithms affecting public discourse. Panels further analyzed the concept of search or internet ‘neutrality’. Data Privacy Law Amendments: Pseudonymization, Research Utilization, Data Transaction At the event, held on May 7, 2020, panels discussed various issues relating to the recent amendments to Korea’s Data Privacy Laws¹. The introduction of the ‘pseudonymization’ concept in the laws was extensively discussed in the broader context of deidentification techniques such as differential privacy. The concept of ‘research utilization’ newly introduced in the laws was also heavily discussed as it defines the scope of utilization of pseudonymized personal information. Panels also discussed the market design of data transaction. Feb. 2020 [Co-hosted with Korean Association for Artificial Intelligence and Law] ‘AI and Data Privacy Laws Amendment’ Seminar Various expectations and concerns coexisted with the “Three Data Act” amendment on January 9, 2020. Panels discussed the expectation that an environment in which data can be better utilized to serve artificial intelligence technology versus the concern that fundamental human rights such as right to privacy may be jeopardized. Panels further discussed the necessary elements of the followup enforcement decrees and the appropriate actualization of the ideals of the amendments. Apr. 2020 ‘Personal Information Governance in the AI Era’ Webinar The Personal Information Protection Act, amended in January 2020, went into

effect in August after a procedure for establishing enforcement decrees. Various expectations and concerns coexisted over the amended law. In this webinar, panels discussed what changes will occur in the governance of personal information, including changes to the regulatory framework due to the amendment. The Personal Information Protection Committee, which newly launched, was given powerful investigation and law enforcement powers. Panels reviewed the law enforcement system that the amended Act envisioned and discussed implications from other existing “committee” experiences and implications from overseas cases.

DAIG (Data & AI Governance) Magazine Publication: In December 2020, SAPI launched the DAIG Magazine, a magazine specializing in data and artificial intelligence legal policy. The first issue was planned with the theme of “Data Governance in the Digital Era”. DAIG was envisioned to serve as a platform for discussion regarding issues of data and artificial intelligence. DAIG wishes to bridge the gap between the academia and the industry. The following articles were published in the DAIG’s first issue. [SAPI: The original articles are in Korean and the English translation of the titles below are provided by SAPI and not confirmed by the authors.] Data Governance in the Digital Era - Data Valuation – Byoung-Pil Kim - Data Collection: Methods and the Current State – Jong Yoon Kim, Haksoo Ko, Byoung-Pil Kim - E-commerce and Personal Information – Doil Son - Use of Medical Data: Issues After the Data Privacy Amendments – Ho Sang Yoon - Collection and Sharing of Financial Data: Changes Under the New Credit Information Act – Joon Young Kim, Bomi Chen, Jee Young Park - Major Issues and Improvement Directions of Data Legislation for AI Industry Development – Hwan Kyoung Ko, Il Shin Lee, Sanghoo Lee Fundamentals of Data and AI Governance - Practical Algorithms – Bonhyo Koo, Byoung-Pil Kim Platform Proliferation - Regulating Online Platforms: Global Trends – Yong Lim, Soojin Lee, Haesung Lee et al. SAPI also held a webinar to commemorate the publication of DAIG on December 22, 2020. Authors presented their papers and held panel discussions in the first half of the event. In the second half, students who participated in the 2020 Summer Research Assistantship Program presented their research. ‘Data Economy’ is a summary of legal, economic, and social issues related to the paradigm shift in the data-driven economy. Various issues such as artificial intelligence, autonomous vehicles, fintech, personal information, internet certificate system, data deidentification, big data and competition laws were reviewed through the lens of data economy, ‘Data ownership’ starts with the recognition that it is difficult to endow ownership of data in a strict legal sense. The book analyzes certain important characteristics of data, and

discusses newly emerging legal issues. The book contains the authors' thoughts on what kind of legal rights can be established for data. Since 2019, SAPI has been publishing issue papers on major research topics regarding artificial intelligence and data.² [SAPI: The original articles are in Korean and the English translation of titles below are provided by SAPI and not confirmed by the authors.

Current Research Activities: Projects: 1 Financial Artificial Intelligence Policy & Ethics Guidelines We investigate use cases of artificial intelligence in the financial sector, discover cases where laws and other regulations related to the financial industry hinder the use of artificial intelligence, and propose regulatory improvement measures. We also review specific policy frameworks to encourage broader adoption of artificial intelligence and propose ethical guidelines for use of AI in financial sector by surveying foreign AI ethics principles. 2 Utilization of Medical Data and Privacy Medicine is an area where artificial intelligence can make a huge difference. Improving data usability is the biggest hurdle for the industry. Privacy consideration based on the special nature of medical data is crucial in balancing data usability and medical privacy. We review issues related to use of medical data and propose a balanced approach. 3 Artificial Intelligence in Court Legal documents including court rulings are mostly in text and thus can be utilized using natural language processing. However, many obstacles exist in implementing AI in judicial proceeding. We survey foreign precedents and recommend policy direction to Korean Courts. 4 Legal Prediction AI We review whether implementation of a legal prediction algorithm in damages claims is feasible. We review and confirm the accuracy of the prediction algorithm. 5 Algorithmic Accountability and Governance We review how the opacity of artificial intelligence leads to controversy over fairness and accountability. In this study we pursue an integrated review of technical, medical, ethical, and legal perspectives. 6 Explainable AI While algorithmic decision-making is proliferating, the difficulty to explain AI's decision hinders trust in the technology. Through this research we survey legal and technical discussions around algorithmic explainability and interpretability. 7 Artificial Intelligence Ethics Principle Decisions by AI may not be fair and may exacerbate social bias. Algorithmic fairness has become a central issue in AI research, both normatively and technically. We survey discussions on algorithmic fairness and propose guidelines grounded in domestic situation. 8 Algorithms and Competition Law here are various perspectives on how algorithmic decision affects market competition. We review and organize various views on AI from the competition law perspective. 9 De-identification and

Pseudonymization of Personal Information With the amendment of Korean data privacy laws, pseudonymization of personal information is a hot research topic involving various technical and procedural challenges. We review practical pseudonymization techniques and propose policy design.

10 Legal DB and AI Utilization We are building pilot legal database to confirm possibility of implementing AI prediction models in judicial proceedings.

45. Note on the Use of Generative AI in Education at Kyushu University

In the "Basic Stance on the Use of Generative AI in Education at Kyushu University" (hereinafter referred to as the "Basic Stance on Generative AI") published in August 2023, Kyushu University has established a basic stance on the use of generative AI in education. In the "Basic Stance on the Utilization of Generative AI in Education at Kyushu University" (hereinafter referred to as "Basic Stance on Generative AI") released in August 2023, the University states that it will provide the necessary education so that students will be willing to learn about new technologies such as Generative AI, correctly understand their mechanisms, origins, characteristics, and limitations, and will have the qualities and abilities to utilize them appropriately.

In this direction, the use of generative AI is basically recommended in the education of the University. On the other hand, it is also true that there are some points that require caution, such as the possibility of violating laws and regulations depending on how generative AI is used.

For this reason, while taking into account the basic stance of generative AI, the following points of caution for the use of generative AI for faculty members in education at the University are presented. For students, a separate document summarizing the points to be considered will be made available, so please refer to that document as appropriate.

In this document, "Generative AI" refers to "AI that generates sentences, images, etc. in response to questions, work instructions (prompt input), etc.". As of September 2023, services such as ChatGPT, Bing AI, Bard, Stable Diffusion, and Midjourney are available.

Please note that these precautions are current as of September 2023, and may be revised in the future in light of changes in social conditions surrounding AI and its operation within the university.

(Rules and precautions for the use of generative AI in education)

(1) Consideration of the possibility of utilizing AI to generate

In deciding whether or not to use generated AI, the decision will be made after clarifying how the use of generated AI or not allowing students to use generated AI will contribute to their learning and personal growth, in light of the basic stance of generated AI, the educational objectives of degree programs and other programs, and the objectives and achievement goals of each class.

In doing so, the possibility of use should be considered on a case-by-case basis, rather than uniformly judging whether or not to use it for the entire class. For example, effective use can be expected in situations such as brainstorming, identifying issues, and verifying the programming code created by the student.

(2) Clarification of specific rules and penalties for use in the class

(1) Whether or not to allow the use of generated AI decided in (1) should be clearly stated in the syllabus, etc. so that students are aware of it. In cases where the use of AI is permitted, the syllabus should clearly state the specific purpose, scope, and method of use, rather than simply stating so in the syllabus, etc.

When clearly indicated in the syllabus, please also refer to the attached typology as appropriate.

If penalties for violating the rules of use are to be established, reference should be made to university regulations, etc., and such penalties should also be clearly stated in the syllabus, etc.

[<Attachment>Rules for using AI to be described in syllabus](#)

(3) Accessibility considerations for generated AI services

When using AI generation tools in class, try to select tools that students can access free of charge, as some services are only available in paid versions. When using paid AI generation tools, efforts should be made to reduce the cost burden on students.

In addition, we will strive to ensure fairness by establishing grading criteria that do not directly affect the quality or differences in the output of the tools used to determine the grade.

(4) Fostering preparedness for the use of generated AI

Please be advised that it is you who decide whether or not to use the output of the generated AI in your submission, and that if you decide to use the output of the generated AI in your submission, you will be fully responsible for the entire content of your submission, even if you did not create it by your own hands. You will be fully instructed that you are responsible for the entire content of your submission, even if you did not create it yourself.

We also remind students that the essence and significance of learning at universities lies in learning on one's own initiative, and that creating reports and other deliverables without one's own initiative, such as using the output of generated AI as-is, does not deepen one's own learning.

(5) Protection of personal and confidential information

When utilizing the generated AI in the classroom, care should be taken to ensure that it does not contain personal or sensitive information. The same care should be taken with data input by students, and appropriate guidance should be given.

(6) Respect for originality and copyright

When allowing students to use the generated AI in their report writing, etc., the method of questions should be devised to encourage students to combine the output results with their own original thoughts and ideas, rather than using the results as is.

In addition, remind students in a timely manner that the output results of the generated AI may infringe on copyright depending on its content and usage.

(7) Explicitly state the use of generated AI

If students are allowed to use the generated AI, they should be instructed to clearly indicate the part of the report in which they cite the output of the generated AI and the name and version of the generated AI service when they use the generated AI in their reports or other documents. Depending on the contents of the class, students may be asked to submit instructions (prompts) to the generated AI and the results of output in response to them (or a series of processes in the case of multiple interactions).

(8) Ensure accuracy of information

Recognize that the user is responsible for checking the accuracy and reliability of the output information without relying on the output of the generated AI, and make sure that students fully understand this.

Guidance should also be provided bearing in mind that the output of the generated AI may not always reflect the most up-to-date information and may include social and cultural bias and discrimination.

(9) Devices in the preparation of assignments and examination questions and in the grading of grades

Consideration will be given to making class assignments and exam questions in a form that requires critical thinking and individual interpretation by the examinee, avoiding as much as possible a form that can be easily solved by a generative AI.

When grading, consider using methods (e.g., oral examinations) that accurately assess how much each student understands on an individual basis, in addition to reports and other artifacts.

(10) Possibility of false detection by AI writing detection software

As of September 2023, there is no way to reliably detect whether a sentence was created by an AI. If AI writing is suspected, the student should be given an opportunity to explain the situation and a detailed investigation should be conducted before a decision is made.

(Other)

It is recommended that teachers receive regular training and incorporate up-to-date knowledge on the evolution of generative AI technology, the social conditions surrounding generative AI, and how to effectively utilize generative AI in education. Please also refer to the "Reference Materials" that follow, as appropriate.

46. About the Use of Generative Artificial Intelligence (ChatGPT, etc.)

About the Use of Generative Artificial Intelligence (ChatGPT, etc.)

Intellectual resilience and flexible sensitivity

Waseda University has been asking students to train themselves to develop “intellectual resilience” so that they can think thoroughly about problems with no correct answer and come up with their own answers, and to foster “flexible sensitivity” that allows them to work together with a variety of people of different races, nationalities, religions, ethnic groups, languages, genders, and sexual orientations in order to solve various problems with respect and sympathy for one another. In today’s world, where we face major issues such as pandemics, wars, and progressing technological innovation, which can be said to be turning points in society, the importance of “intellectual resilience” and “flexible sensitivity” is increasing.

For example, the development of new technology will change the way we live, work, and the state of society. It’s an unavoidable reality that scientific progress and technological

innovation will always have positive and negative sides, advantages and disadvantages. It depends on how humans use technology, and on national policies and institutions that promote or regulate these technological innovations. For those of you who aim to strengthen your “intellectual resilience”, nurture your “flexible sensitivity”, and contribute to society, realizing the correct use of new technology is an extremely important issue. It will also be an important task to conduct research on measures to mitigate the social impact and disseminate the results, with a “flexible sensitivity” that sympathizes with those who are negatively affected by new technology.

Fundamental Attitudes towards Generative AI

Generative artificial intelligence (hereinafter referred to as generative AI) such as ChatGPT (OpenAI), Bard (Google), and Bing (Microsoft), which have become a hot topic in recent years, are advancing their performance at an accelerated pace. It has been pointed out that they have the potential to bring about major changes in our society. For this reason, interactive automated response services based on generative AI are not only tested around the world but also discussed at educational institutions as to how they should be utilized and regulated. Those who advocate “intellectual resilience” need to be familiar with the correct usage of generative AI. To that end, we must have sufficient knowledge of generative AI and cultivate an attitude of keeping pace with technological progress and continuing to innovate our knowledge. Because then you can masterfully use the technology instead of being the one used by it.

In addition, it is imperative to have an accurate understanding of the impact that generative AI will have on society and people, and the issues that arise around the formation process, when we consider the issues of whether to stop the development or regulate the use of generative AI, all of which are in the midst of debate. This is because in the development of new technologies and the production of products and services based on them, a great impact on various people is observed not only after development and production but also before.

Characteristics of Generative AI

Generative AI learns by itself based on the vast amount of data on the Internet, and can

instantly generate response sentences in a natural language dialogue format for various questions and requests in a wide range of fields. Also, even if you ask the same question, the answer (sentence) that is returned will change each time. The following tasks that were previously performed by humans can now be handled by generative AI.

- (1) Generation of new sentences based on given strings and keywords (creation of reports, impressions, etc.)
- (2) Generating answers for tests (fill-in-the-blank questions, multiple-choice questions, open-ended questions, etc.)
- (3) Brainstorming, creation of ideas
- (4) Summarization, translation and proofreading of existing texts and web pages
- (5) Data analysis
- (6) Easy programming
- (7) Lyrics, composition
- (8) Searching for documents and information

If you can use generative AI with these functions well, you will be able to save a lot of your work time, and many people will get the benefit of being able to concentrate on other important tasks.

On the other hand, the following flaws and difficulties have also been pointed out. There are many things that students need to be careful about.

First, generative AI does not always output accurate information. Currently, for example, there are quite a few cases where answers are generated using non-existent universities or companies, incorrect numerical data, or wrong names. It has also been observed often to fabricate bibliographies of titles and author names that do not exist.

Second, it has been pointed out that reports generated by generative AI are (currently) logically correct as text, but the content is shallow and there is a tendency for a repetition of plausible and obvious things. If you use the results of the generative AI without examining the content, the report will be lacking in originality.

Third, if you submit the thesis created by the generative AI without carefully examining it, the thesis will be evaluated as it is, even if there are incompleteness or fatal errors in the content. You may also be subject to disciplinary action if plagiarism, inappropriate citation, or fabrication is confirmed in the submitted thesis, even if you did not intend to do so. Of course, if you submit a thesis created by the generative AI as it is, it will be punished as a fraudulent act similar to cheating.

Fourth, careless use of generative AI may lead to the leakage of confidential information within the university or company, and questions that include your name, etc. may lead to the leakage of personal information. It has also been pointed out that the disorderly use of generative AI can lead to discrimination, human rights violations, and criminal behavior. Freedom of action must be guaranteed for technological progress, but no one has the freedom to infringe on human rights. “Intellectual resilience” must also pay attention to the tension between technology and society.

Usage and Limitations of Generative AI

Surely you might be able to complete a report quickly with generative AI, but you’re giving up the opportunity to improve your abilities. Reports are created by gathering a large amount of information from documents and materials, classifying, analyzing, and summarizing them to create the most appropriate answer to the task. Writing a report is the most important exercise in developing “intellectual resilience”. As a result of neglecting this task and prioritizing short-term profits, in serious situations such as entrance exams and work in the real world, you will notice the fact that you have learned nothing.

On the other hand, for those working as business creators who envision unprecedented products and services, the use of generative AI will lead to the realization of effective work. In the process of collecting, classifying, summarizing, and analyzing vast amounts of information about similar past cases to generate new ideas, the partial use of generative AI could save time, and you could spend more of your time generating important new ideas. This effective use assumes that you know firsthand how much work is needed and what quality is required at

each stage in the process of making a report. That way, you can easily determine the merits of substituting that part with generative AI, as well as discrepancies and mistakes in the responses shown by AI, and you can use it in a suitable manner. It is important to train your “intellectual resilience” daily so that you can use generative AI appropriately according to the time and situation.

In addition, there are reservations about generative AI in terms of “flexible sensitivity”. To be able to work together to solve social problems while respecting the diverse attributes and backgrounds of various people, it is necessary to build mutual trust. It requires both mutual understanding and sympathy. A certain kind of expression is also required in exchange of messages, but it seems difficult for generative AI to appropriately provide expressions that show fine-grained correspondence and persuasive power.

Future Action

You will continue to utilize a variety of products and services developed based on new technologies. Before using them, or while using them, you should repeatedly think about whether they are necessary for you, whether they are beneficial to you, and what kind of impact they have on society. What is beneficial to you now may only have a negative effect on your future self. Some people get more benefits than disadvantages, others do not. Think objectively and with an altruistic spirit. Waseda University’s students, faculty, and staff, as well as all concerned parties, should not forget this way of thinking.

“Intellectual resilience” applies to all matters in the world. To fully understand the multi-faceted impact of new technologies, how to curb the negative aspects, and to further expand the positive aspects, we (not only individuals, but also governments, industries, communities, etc.) must think carefully about what we should do.

47. A Guide to the Use of Generative AI

Preface

Thanks to the advancements in artificial intelligence technology, generative AI can now perform tasks once thought exclusive to humans. This includes ChatGPT, which has recently garnered significant attention. Generative AI can solve complex problems, create works of art, and both understand and generate language. These amazing abilities of generative AI make us rethink about originality and creativity, which we have long regarded as one of the core values of research and education. 'Does using generative AI in research undermine the originality of the research?' 'Will using generative AI in education hinder the development of students' learning and creativity?' It is hard to give a clear answer to these questions. However, it is certain that the advent of generative AI will somehow influence our thinking and learning methods. The UNIST Education Innovation Task Force has recognized these issues and conducted a survey of faculty, students, and researchers to understand the impact of generative AI on their education, learning, and research activities. We also considered ways to minimize side effects and effectively utilize Generative AI. The result of these deliberations is this guidebook. This guidebook includes a brief introduction to generative AI technology (pp.4-9), guidelines for using generative AI for faculty, students, and researchers (pp.10-15), and tips for using generative AI (pp.17-40). Also, the appendix (pp.43-50) includes the results of a survey on the usage and perception of generative AI targeted at UNIST's faculty, students, and researchers. Artificial intelligence technology is progressing at a tremendous speed. We must continue to analyze the impact of rapidly changing artificial intelligence technology on research and educational activities, and take appropriate action in response to any ensuing challenges. The UNIST Education Innovation Task Force will continue to make efforts to develop policies and guidelines necessary to utilize the advantages of artificial intelligence while minimizing side effects and maintaining and improving the quality of research and education.

What is Generative AI?

The Diversity of Generative AI “Generative artificial intelligence or generative AI is artificial intelligence capable of generating text, images, or other media in response to prompts” (“Generative artificial intelligence,” 2023). Generative AI can be classified in various ways, depending on what it creates. • Text Generative AI: This type of AI uses natural language

processing technology to generate text. It can write novels, news articles, scripts, and poetry, and generate answers to specific questions. Notable examples include ChatGPT, Bing, and Bard.

- Image Generative AI: Image generative AI can generate images according to given descriptions or improve existing images. DALL-E, Midjourney, and Stable Diffusion are some representative examples.
- Voice Generative AI: Voice generative AI can convert text to speech or transform one voice into another. Examples include Meta Platforms' VoiceBox and Google's Tacotron.
- Music Generative AI: Generative AI can also generate music. Google's MusicLM is a notable example.
- Video Generative AI: Video Generative AI can generate videos based on text or image input, or modify existing videos. Examples include RunwayML's Gen2 and Meta Platforms' "Make-A-Video."

Generative AI vs Discriminative AI

Generative AI produces new content, but it does not create something from nothing. Generative AI learns the distribution of data from the provided input and probabilistically generates new content corresponding to input values. For example, a model that generates text creates sentences that are likely to follow the input sentence, based on the probability distribution of the pre-trained data. To boil it down, generative AI generates new sentences based on learning that certain phrases are likely to follow certain other phrases. Generative AI typically undergoes pre-training based on a vast amount of unlabeled data.¹ In contrast, discriminative AI, unlike generative AI, generally learns from pre-labeled data. It then outputs labels corresponding to input values based on this learning. Such discriminative AI is used for classification or prediction.

Capabilities of Generative AI

Discriminative AI is chiefly utilized for tasks such as classification or prediction. However, generative AI is also capable of executing these tasks. It is both conceptually and technically possible to employ generative AI in these roles. Firstly, from a conceptual standpoint, if we instruct the generative AI to generate classification or prediction results, it's possible to perform these tasks using generative AI. Secondly, while generative AI does not explicitly learn from labeled data for specific classification or prediction tasks, it learns from a vast amount of data and naturally becomes capable of carrying out classification or prediction tasks. As the performance of generative AI improves, its ability to handle a more diverse and comprehensive range of tasks also increases. According to Narang and Chowdhery (2022) of Google Research, “as the scale of the model increases, the performance improves across tasks while also unlocking new capabilities,” as follows:

The emergence of transformer technology has

made it possible for generative AI models to scale in such a way.² Thanks to this transformer technology, we have been able to handle vast amounts of data in the field of Natural Language Processing (NLP), leading to the emergence of Large Language Models (LLMs).³ These LLMs, trained on a massive scale of data, have begun to demonstrate a range of abilities beyond simple language processing. Now, generative AI trained on a massive amount of data can perform various tasks such as not only translating text, but also logical reasoning, problem-solving, summarizing text, writing code, recognizing patterns, and composing text. With the ability to process language, it has become possible to associate text with various types of data, such as images, voices, and videos, for combined learning. Artificial intelligence capable of understanding and processing various media types is referred to as multimodal generative AI. Although not yet publicly released, OpenAI's GPT-4 is reportedly capable of understanding photos or graphs and of providing logical reasoning or problemsolving capabilities (OpenAI, 2023). Artificial intelligence technology continues to evolve. With the increase in computing power and the improvement of AI algorithms, generative AI is gradually approaching Artificial General Intelligence (AGI), which has the potential to perform any intellectual task that a human can do.

Limitations of Generative AI Generative AI still has its limitations. A prime example is 'hallucination.' Hallucination refers to the phenomenon where generative AI provides answers that are not factual, treating them as if they were true. This occurrence of hallucination arises because generative AI does not generate the absolute correct answer, but instead produces the response that is most likely to be correct. Various methods are being explored to overcome this phenomenon of hallucination. These include approaches like post-processing, which filters the generated answers, or fine-tuning to enhance the model. Additionally, augmenting the volume and quality of the training data used by generative AI can help to mitigate the occurrence of hallucinations. However, completely overcoming the issue of hallucination in generative AI is a significant challenge. As a result, it is crucial for users to scrutinize and validate the responses from generative AI carefully. Another limitation lies in the potential for generative AI to generate biased content. Since generative AI learns patterns from its training data, biases or inclinations present in that data may be mirrored in the model. As a result, the responses produced by the AI could exhibit bias or unfairness. This means biases concerning specific races, genders, or societal issues could be reflected, posing substantial ethical concerns. Therefore, it's crucial for users to maintain a critical perspective.

Guidelines on Using Generative AI For Faculty

Education Try using generative AI yourself: Whether you decide to allow generative AI in your course, it is helpful to test them to understand how they respond to your assignments and how students might use them. This will give you insight into the strengths and limitations of Generative AI, and you may discover ways to incorporate them into your teaching. Make a course-level decision on the use of generative AI: There is no institution-wide prohibition on using generative AI for teaching and learning purposes. You may choose to incorporate them into your course or prohibit them. The decision to utilize generative AI should be made at the course level autonomously by the instructor, based on the characteristics of each course. Update the syllabus: Include your course policy regarding the use of generative AI in the syllabus, and make sure to announce it. Use various forms of assessment: Evaluating students based solely on one final project or assignment can increase the possibility of academic dishonesty. It can be beneficial to allocate scores for a sequence of tasks such as proposing ideas, improving them, and reflecting on their learning. Furthermore, oral exams or unexpected questions can be effective ways to assess whether students thoroughly understand their submitted work and if they have completed it on their own. Require citation in assignments: Merely requiring students to strictly cite sources in their assignments prompts them to be more cautious when using generative AI. Consider using an AI detector: If you are concerned about students submitting essays generated by AI tools, you can use AI detectors to identify them. However, it's important to note that identifications should not rely solely on these detection tools. Content should still be carefully reviewed for the final decision. Discerning the use of generative AI can be challenging

Research Check the policy on AI tools of the academic journal or conference: Academic journals or conferences may have different policies regarding the use of generative AI. It's necessary to check in advance if there are specific guidelines about the use of generative AI in the journal you plan to submit to or the conference you plan to participate in. Cite generative AI tools: Cite generative AI tools to acknowledge their contribution to your research. Verify the source: Generative AI learns from existing data to generate new content. When using generative AI in research, it's crucial to conduct additional searches to ensure that the output doesn't overlap with existing work or violate any copyrights. Be mindful of bias in generative AI: Information in the

training data can carry inherent biases, which may then be reflected in generative AI models. This can inadvertently result in biases related to specific races, genders, or social issues being incorporated into your research. Do fact-check: Generative AI may produce fake or incorrect content. You should fact-check any information obtained from Generative AIs before using it.

Security Do not share important research information with AI tools: Many generative AI tools are open platforms, and third parties may record or analyze conversations with AI. Avoid sharing any confidential or proprietary research information with AI that could potentially be exposed to third parties. Ensure your devices have up-to-date security software: Ensure your devices have up-to-date security software and firewalls to prevent unauthorized access by Generative AI to sensitive information. Do not share personal information with Generative AI: Avoid sharing sensitive information such as your name, address, phone number, or other personal details when interacting with Generative AI. Be cautious when using AI-generated code: The code produced by AI may contain errors or security vulnerabilities. It's crucial to go through a verification process before utilizing AI-generated code.

Guidelines on Using Generative AI For Students

Learning Be mindful of the course policy: Your professor may prohibit using generative AI. Carefully read the current syllabus and follow the course policy. Using generative AI could be considered academic misconduct. Do fact-check: Generative AI may produce fake or incorrect content. You should fact-check any information obtained from Generative AI before using it. Think critically: You should be aware of the drawbacks and limitations of the generative AI you are using. Always question and critically evaluate the AI's output. Furthermore, you should strive for results that surpass those generated by the AI. Use generative AI as an auxiliary tool only: The creativity and problemsolving skills gained through personal experiences cannot be replaced by generative AI. Relying on generative AI could hinder your learning and growth

Research Check the policy on AI tools of the academic journal or conference: Academic journals or conferences may have different policies regarding the use of generative AI. It's necessary to check in advance if there are specific guidelines about the use of generative AI in the journal you plan to submit to or the conference you plan to participate in. Cite generative AI tools: Cite generative AI tools to acknowledge their contribution to your research. Verify the source:

Generative AI learns from existing data to generate new content. When using generative AI in research, it's crucial to conduct additional searches to ensure that the output doesn't overlap with existing work or violate any copyrights. Be mindful of bias in generative AI: Information in the training data can carry inherent biases, which may then be reflected in generative AI models. This can inadvertently result in biases related to specific races, genders, or social issues being incorporated into your research.

Security Do not share important research information with AI tools: Many generative AI tools are open platforms, and third parties may record or analyze conversations with AI. Avoid sharing any confidential or proprietary research information with AI that could potentially be exposed to third parties. Ensure your devices have up-to-date security software: Ensure your devices have up-to-date security software and firewalls to prevent unauthorized access by Generative AI to sensitive information. Do not share personal information with Generative AI: Avoid sharing sensitive information such as your name, address, phone number, or other personal details when interacting with Generative AI. Be cautious when using AI-generated code: The code produced by AI may contain errors or security vulnerabilities. It's crucial to go through a verification process before utilizing AI-generated code.

Guidelines on Using Generative AI For Researchers

Research Check the policy on AI tools of the academic journal or conference: Academic journals or conferences may have different policies regarding the use of generative AI. It's necessary to check in advance if there are specific guidelines about the use of generative AI in the journal you plan to submit to or the conference you plan to participate in. Cite generative AI tools: Cite generative AI tools to acknowledge their contribution to your research. Verify the source: Generative AI learns from existing data to generate new content. When using generative AI in research, it's crucial to conduct additional searches to ensure that the output doesn't overlap with existing work or violate any copyrights. Be mindful of bias in generative AI: Information in the training data can carry inherent biases, which may then be reflected in generative AI models. This can inadvertently result in biases related to specific races, genders, or social issues being incorporated into your research. Do fact-check: Generative AI may produce fake or incorrect content. You should fact-check any information obtained from generative AI before using it

Security Do not share important research information with AI tools: Many generative AI tools are open platforms, and third parties may record or analyze conversations with AI. Avoid sharing any confidential or proprietary research information with AI that could potentially be exposed to third parties. Ensure your devices have up-to-date security software: Ensure your devices have up-to-date security software and firewalls to prevent unauthorized access by Generative AI to sensitive information. Do not share personal information with Generative AI: Avoid sharing sensitive information such as your name, address, phone number, or other personal details when interacting with Generative AI. Be cautious when using AI-generated code: The code produced by AI may contain errors or security vulnerabilities. It's crucial to go through a verification process before utilizing AI-generated code.

Tips for Using Generative AI

Education (Faculty) Generative AI can be a valuable tool for educators. For example, it can make class preparation more efficient. By simulating student responses with generative AI, educators may be able to predict student reactions, potentially maximizing the effectiveness of instruction. However, in classes where the use of generative AI by students needs to be limited, it is crucial to design the course in accordance with the specific characteristics of the subject matter. Course Design

1. Try asking generative AI for course-specific integration advice. It can also be helpful to directly ask generative AI how it can be utilized in education according to the characteristics of each subject. Explain what subject you are teaching and ask how you can use generative AI in that specific subject.
2. Try using the flipped learning strategy. If you are concerned about students using generative AI for assignments, adopting a teaching method known as flipped learning or the flipped classroom could be a good strategy. “A flipped classroom is an instructional strategy and a type of blended learning, which aims to increase student engagement and learning by having pupils complete readings at home and work on live problem-solving during class time.” (“Flipped classroom”, 2023) In a flipped classroom environment, students learn independently based on the course materials provided before class and perform tasks such as assignments and discussions during class time when the use of generative AI can be regulated. This approach can naturally alleviate concerns

about misconduct caused by the use of generative AI. In the flipped classroom, it's crucial to thoroughly evaluate various activities that occur during class, such as student presentations and questions. If these diverse activities are included in the assessment criteria, students are likely to participate more actively in class.

3. Try asking the generative AI for textbook recommendations for your course. If you're developing a new course, you might want to get recommendations for textbooks or reading materials from the generative AI. GPT-3.5 and GPT-4 were trained based on data up until September 2021. If you require more recent information, consider using Google's Bard or Microsoft's Bing.

4. Consider using generative AI in syllabus creation. After explaining the course to the generative AI, ask it to draft an initial syllabus. If there are specific contents that need to be included in the syllabus, request the AI to incorporate them.

Class Preparation 1. Try using generative AI to write your lecture outlines. If you want to ensure your lecture proceeds in an organized manner, having an outline ready in advance can be helpful. After explaining the content you will cover in class to the generative AI, try delegating the task of drafting the lecture outline. Once the outline is created, you can also create a draft of the presentation slides through generative AI. Try asking the generative AI to convert the outline into HTML. Next, save the HTML code in Windows Notepad, then import it into Office365's Word. From there, exporting as a PowerPoint presentation will quickly complete a draft of the PowerPoint slides.⁴

2. Try asking the generative AI to generate explanations in various ways. Depending on the students' levels, different types of explanations may be required. Consider requesting a variety of explanations from the generative AI.

3. Try using generative AI to create various examples that explain concepts. When explaining concepts to students, using various examples can greatly aid their understanding. However, finding these examples can be time-consuming. Consider asking the generative AI to create examples.

4. Try using generative AI to simulate potential questions from students. Inform the AI about the topic for the next class and ask it to predict possible questions. This could enable you to respond to students' queries more effectively.

Assessment 1. Consider using oral exams. If you are concerned about the potential decline in students' academic performance due to generative AI, implementing oral exams or unexpected

questions can be an effective method to evaluate students' comprehension of their submitted assignments. Do not let the learning process stop at assignment submission; instead, encourage students to articulate their understanding of their own work. This approach can prevent students from submitting assignments they have not fully grasped.

2. Try asking generative AI to create quiz problems. You can ask the generative AI to form True/False or multiple-choice questions. By revising and utilizing the questions generated by the AI, you can save time on creating quiz problems.

3. Consider Using AI Detectors. If you prohibit students from using AI generators to write and submit essays, it is recommended to use detectors that reveal whether generative AI was used. <https://www.turnitin.com/> <https://gptzero.me/> The Education Innovation Task Force tested several AI detectors launched by May 2023 and found that Turnitin and GPTZero showed relatively superior performance. As a result, an AI detection functionality was added to the Turnitin service, which was previously used to check the similarity of papers and assignments. Turnitin is integrated with our university's Learning Management System (LMS), Blackboard. Encourage your students to submit their assignments via Blackboard. The setup for Turnitin in Blackboard is as follows: When using Turnitin, the percentage of AI usage is displayed in the blue box at the bottom right of the report. While AI detectors are valuable tools, they should not be wholly relied upon. An AI detector is just one reference tool. If there's a suspicion that a student's work was generated by AI, it could be helpful to employ multiple AI detectors like GPTZero, where possible. For shorter texts, the accuracy of AI detectors can decrease. Consequently, it's advisable to use AI detectors on longer texts, when possible. Determining whether a student has used a generative AI to ghostwrite their assignment can be challenging. For an appropriate assessment, it's recommended to meticulously review the content of the student's work, evaluate their level of understanding through discussions, and then draw a conclusion about the potential use of AI.

Grading and Feedback

1. Try using generative AI for grading. Generative AI can serve as a grading assistant. However, when using generative AI for grading, it is essential for the instructor to personally review and adjust the student's essay and grading results. Also, please notify the students that generative AI was used in the grading process. Start by creating a grading rubric, as shown in the example below.

5 Generative AI typically recognizes copied and pasted table content quite well. However, to ensure it is well-understood by the generative AI, it is recommended to modify it into a more comprehensible format. This process can also be facilitated by the generative

AI. For example, you can input “Reorganize the following essay rubric:”, followed by the simply copied rubric table. Then, a reorganized grading rubric will appear as follows: After double-checking the content, ask the generative AI to grade the essay based on that rubric: Then, the generative AI will provide the grading results as shown below:

7. Try using generative AI for feedback. Besides grading, generative AI can also assist in crafting initial feedback for students. Ask the generative AI to generate feedback in a way that encourages your students. In using generative AI to give feedback to students, it is crucial that you always check the content of the feedback.

Education (Students) Consider utilizing generative AI in your learning process. AI can provide a customized learning experience tailored to your individual level and pace, and it can assist in deepening your understanding across a range of topics. However, it is crucial to understand the limitations of generative AI and ensure that you do not miss opportunities to develop your own competencies.

1. Consider developing a study plan using generative AI. You can ask the AI to help create a study schedule either for the semester or the holidays. Set your desired field of study and duration, then ask for recommendations such as textbooks, MOOCs, or YouTube content. 2. Request concept explanations from generative AI. Ask the AI about concepts you don't understand. If there are areas that you're finding difficult to comprehend, continue asking questions based on the AI's explanations. Consider asking the same question multiple times or posing it to a different generative AI. Just as you might get a better answer by asking the same question to multiple people, you could find a better answer by asking different AIs the same question. 3. Consider using generative AI to draft an email to your professor. If you have questions but find it challenging to approach your professor directly, sending an email can be a good alternative. Generative AI can provide significant assistance in writing emails. Ask generative AI for advice on the most effective way to draft your email. Also, it is a great way to learn how to compose an email politely and professionally.

Research (Faculty, Students, Researchers) When utilizing generative AI in research, you can enhance your work's efficiency. However, as the reliance on generative AI increases, the originality of your research might gradually decrease. Determining the acceptable boundary for using generative AI is challenging. This is similar to removing a single grain from a heap and asking when it no longer qualifies as a heap. If you're employing generative AI in your research,

it is advisable to use it judiciously, ensuring that it doesn't undermine your work's originality. Always remember, the responsibility for the final outcome of the research lies solely with the researcher.

1. Verify the editorial policy of the journal or publisher. It is essential to check the editorial policies of prospective journals or publishers not only at the submission stage but also throughout the research process. Below are the editorial policies from several journals and publishers. Please note that these policies may change over time

2. Disclose how you used generative AI. It is important to disclose how you used generative AI in your research. Please make it clear in the appropriate places of your paper, such as the acknowledgements or methods, about the use and application of generative AI. Nature's editor-in-chief, Magdalena Skipper, emphasizes, "authors using LLMs in any way while developing a paper should document their use in the methods or acknowledgments sections, if appropriate" (Chris Stoke-Walker, 2023).

3. How to cite generative AI There is a risk of being perceived as committing plagiarism if the usage of generative AI is not appropriately cited in academic writing. Holden Thorp, editor-in-chief of the Science family of journals, has made his position clear on this issue. He states, "use of AI-generated text without proper citation could be considered plagiarism" (Chris Stoke-Walker, 2023). Researchers must appropriately cite the use of generative AI, adhering to the style guidelines of each specific journal. Here are a few methods for appropriately citing generative AI in scholarly papers

4. Try using generative AI for proofreading your paper. You can use generative AI for proofreading your paper as follows. However, excessive editing by AI could alter the content and results of your paper, potentially compromising the originality of your work. Therefore, authors should be cautious when using generative AI for proofreading.

5. Try using generative AI for your literature review. Ask for recommendations of the most frequently cited or most influential papers in a specific field. (Please note that ChatGPT is based on data trained until September 2021. For the latest materials, try using Google's Bard or Microsoft's Bing.) If you are a subscriber of the paid service GPT-plus, you can search the archives of Arxiv, a prominent academic paper repository, through a plugin. Try searching for "arXiv" in the plugin store. The Chrome browser features a variety of plugins related to generative AI. One such plugin, ArxivGPT, can provide crucial information about papers uploaded on Arxiv.

Appendix: Survey Results The Education Innovation Task Force conducted a survey to understand how generative AI is used and perceived by UNIST members. The aim of this survey was to formulate effective strategies for handling generative AI. The survey details are as follows:

Survey

period: April 28, 2023, to May 7, 2023 Participants: 70 faculty members, 147 graduates, 105 undergraduates, and 36 researchers Have you ever used generative AI like ChatGPT? Among the respondents, 89% of the faculty, 93% of graduate students, 90% of undergraduates, and 69% of researchers indicated they have used generative AI. Among the respondents, 43% of graduate students and 65% of undergraduates had used generative AI in their classes. In the case of undergraduates, about two-thirds of the respondents are using generative AI in their courses. How often do you use generative AI? Among those who responded that they have used generative AI in the previous question, 38% of faculty, 51% of graduate students, 39% of undergraduate students, and 16% of researchers reported using it five times a week or more. Graduate students showed the highest frequency of use. What language do you primarily use when using generative AI? Among those who responded that they have used generative AI, 83% of faculty members, 73% of graduate students, 72% of undergraduate students, and 72% of researchers revealed that they primarily use English when leveraging generative AI. The high prevalence of English usage among students seems to be attributable to the fact that courses at UNIST are conducted in English. When using generative AI, are you cautious about the potential leakage of personal information or important research data? Among those who reported using generative AI, 64% of faculty members, 71% of graduate students, 60% of undergraduate students, and 80% of researchers stated that they are cautious about security when using generative AI. What generative AI tool do you mainly use? Among the students who reported having used generative AI, most responded that they use the GPT-3.5 (free) version, while 21% of graduate students and 18% of undergraduates who have used generative AI reported using the paid services GPT-4.0 or GPT (API). Other responses included Bard, Github Copilot, ChatPDF, Stable Diffusion, and Perplexity.ai. Generative AI can produce answers that are different from the truth. Do you fact-check the answers obtained through it? 57% of graduate students and 39% of undergraduates who have used generative AI responded that they always fact-check the output. Only 1% of graduate students and 3% of undergraduates reported that they do not fact-check at all. These results suggest that students are well aware of the limitations of generative AI If you have used generative AI for coursework, specifically for what purposes have you used it? (Multiple responses) 81% of graduate students and 75% of undergraduates who reported using generative AI in their coursework indicated that they use it for assignments. Both graduate and undergraduate students were found to use generative AI more for reviewing than for previewing. 16% of graduate students and 13% of undergraduates reported

using generative AI for exams. If you have used generative AI for coursework, how specifically have you used it? (Multiple responses) The application of generative AI revealed a very similar pattern among both graduate and undergraduate students. Both groups primarily use it for coding. When using generative AI, only 3% of both graduate and undergraduate students said they submit their writing assignments without any revisions. Do you think using generative AI is helpful for your learning? In a survey conducted among students who use generative AI in their studies, the majority indicated that the use of generative AI helped or significantly aided their learning. However, 3% of undergraduates responded that it seemed to obstruct their learning. Of the faculty respondents, 18% indicated they plan to ban the use of generative AI in future assignments or exams. However, 63% reported they would not prohibit the use of generative AI in these areas. Additionally, 52% of faculty expressed willingness to incorporate generative AI into their future classes. Overall, this data suggests a generally favorable attitude towards generative AI usage among faculty. In response to the question of whether they would be willing to use an AI detector assuming that the use of generative AI by students is banned, 49% of respondents answered affirmatively. Furthermore, 56% of faculty members indicated that changes in the assignment or exam formats are necessary due to generative AI. How do you utilize generative AI in your research? (multiple responses) For faculty members, graduate students, and researchers, the most common use of generative AI in research was for translation or proofreading. In contrast, undergraduate students used generative AI for brainstorming and coding more frequently, with 56% and 59% respectively, compared to other groups. Plagiarism is generally defined as "taking someone else's ideas or creations without proper citation." Do you think using generative AI for writing papers constitutes plagiarism? Two-thirds of the respondents, consisting of 66% faculty members and 67% researchers, believe modifications are needed to the existing definition. Additionally, 42% of graduates and 54% of undergraduates also responded that they think the current definition needs revision. Which of the following do you think considered cheating when students use generative AI for their assignments? (Multiple responses) While 67% of faculty members responded that not citing the use of generative AI constitutes cheating, only 39% of graduate students and 48% of undergraduates shared this view, demonstrating a disparity in perception. Furthermore, the percentage of faculty members who believe that submitting work without modification constitutes academic dishonesty was 56%, lower than that of the students.

48. SUM Framework for the use of Generative AI Tools

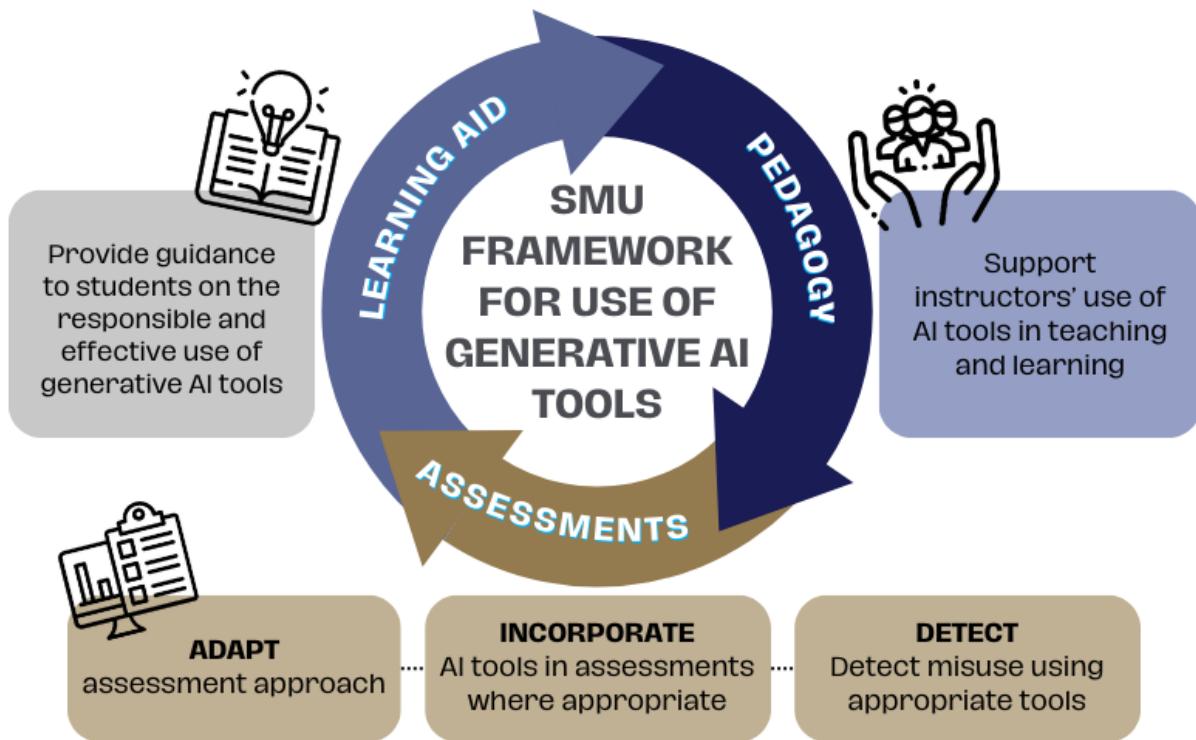
INTRODUCTION

Singapore Management University has developed a **Framework for The Use of Generative Artificial Intelligence (AI) Tools** which aims to provide guidance on how the SMU community can use such tools, and where resources should be channeled to ensure the responsible and effective use of generative AI tools. It provides the University's position on the use of generative AI tools, as well as its use in the following specific contexts: [i] as a learning aid for students or [ii] as a pedagogical tool for instructors, and [iii] in assessments.

UNIVERSITY'S POSITION ON THE USE OF GENERATIVE AI TOOLS

The University recognises the benefits that generative AI tools bring and its increasing presence in our everyday lives. Our students will build their careers in a world where AI tools will be increasingly pervasive and it is incumbent on us to teach students how to interact with such tools in an effective and discerning way. With fast evolving technologies like AI, our response is not to shy away from it, but to assess how we can harness the best of what it can offer and adapt the way we teach and approach assessments to prevent misuse. We are committed to exploring ways to integrate generative AI tools into education, but at the same time safeguard academic integrity and standards of academic rigor. For this reason, ensuring the responsible and ethical use of generative AI tools must be part of our response.

THE FRAMEWORK



SMU's position

The University recognises the benefits of generative AI tools and is committed to integrating generative AI tools into education, while safeguarding academic integrity and rigour. At the same time, the University aims to teach students to use these tools responsibly and effectively.

USE OF GENERATIVE AI TOOLS BY STUDENTS AS A LEARNING AID

Generative AI tools have the potential to enhance students' understanding of knowledge leading to better learning outcomes. Nonetheless, students need to be made aware of the limitations and ethical use of such tools.

The Student Success Centre has launched the **Guide to Learning with AI** online module to educate students on how generative AI tools can be used as a learning aid in an effective and responsible manner. This module can be found on eLearn.

USE OF GENERATIVE AI TOOLS BY INSTRUCTORS AS A PEDAGOGICAL TOOL

Generative AI tools can lead to improved teaching quality and support instructors to create educational content.

The Centre for Teaching Excellence (CTE) has developed a resource page on the "["Use of AI Tools in Assessment and Teaching"](#)" for instructors and will be updating it as new information on AI tools arises. In addition, a [series of webinars / workshops](#) has been planned to provide support to

instructors on the effective use of generative AI tools in teaching, including demonstrating best practices for incorporating these tools in their teaching.

USE OF GENERATIVE AI TOOLS BY STUDENTS ON ASSESSMENTS

The University will address the potential misuse of generative AI tools in assessments using the following three-pronged approach of **Adapt**, **Incorporate**, and **Detect**:

1. **ADAPT** by changing assessment approach

Instructors can adapt by redesigning assessment questions such that they are beyond the existing capabilities of generative AI tools or changing the type of assessments employed.

2. **INCORPORATE generative AI tools on assessments where appropriate.**

Instructors may choose to allow the use of generative AI tools on assessments, where its use can enhance learning. If an instructor chooses to allow the use of a generative AI tool:

- its permitted use must be made explicit to students;
- guidelines for when and how it is to be used should be provided;
- use of the tool should be credited using a widely accepted format, where applicable, such as the APA Style Guide or Chicago Manual of Style.

3. **DETECT misuse of generative AI tools with detection tools, as they become available.**

SMU instructors should follow the “**DRIVE**” approach if they encounter suspected cases of students’ unauthorised use of generative AI tools during assessment. Please refer to the section below for more details on the **DRIVE approach**.

Communicating Expectations to Students

Instructors should clearly communicate their expectations and policies regarding the use of Generative AI tools and ensure that students understand the importance of academic honesty and originality. For assessments that do not allow students to use generative AI tools, students should be informed of the consequences of violation.

For assessments where the use of generative AI tools is permitted, students should be told the specific areas and tasks where generative AI tools could be applied and what constitutes unauthorised use.

In this link ([Statement Templates](#)), instructors can access templates of statements they can adapt according to their requirements:

1. Statement to explain the use of Generative AI tools in assessment to students.
2. Statements of Declaration for students in assessments where the use of Generative AI is permitted, restricted, partially permitted under certain conditions specified by the instructor.

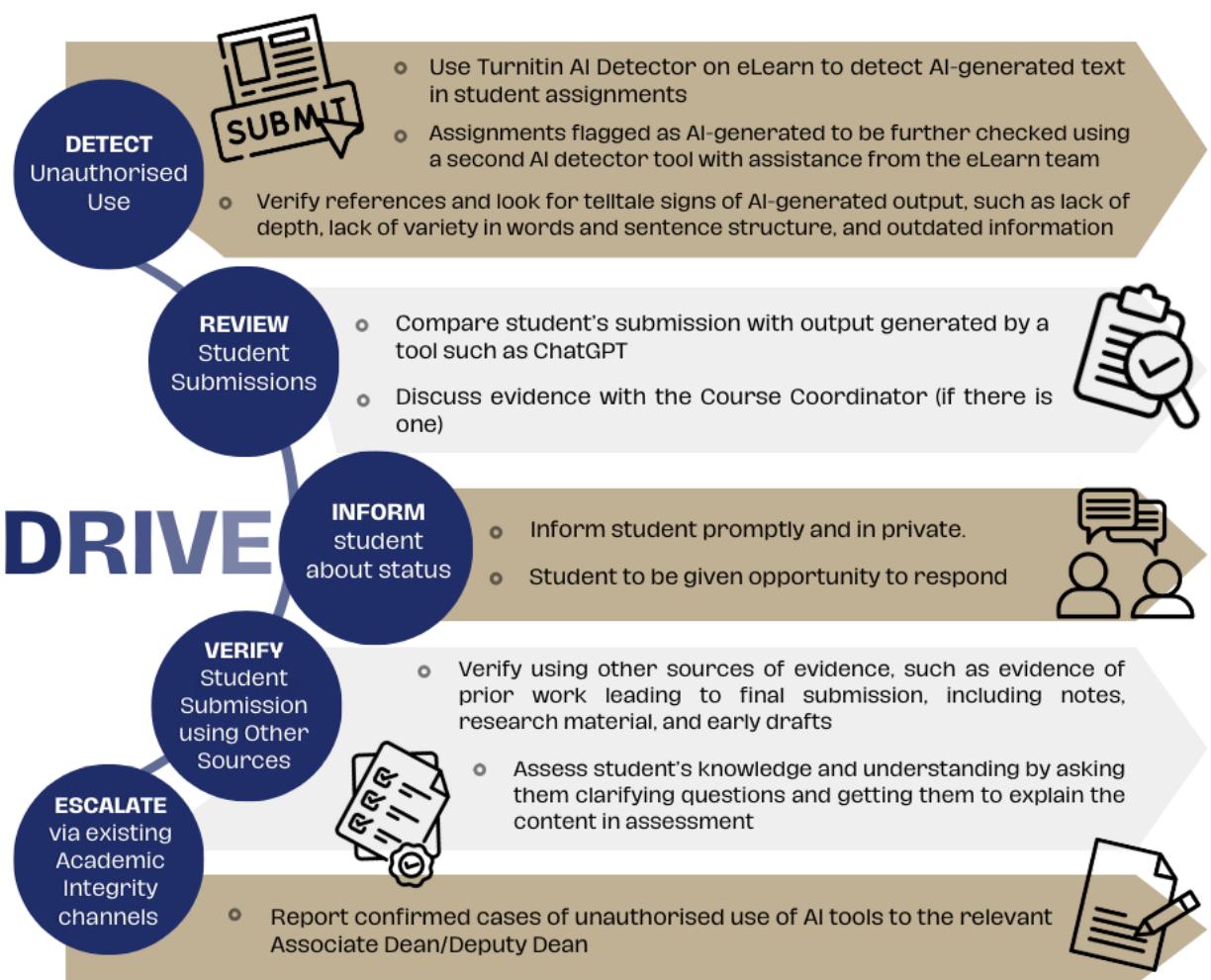
In the event suspected unauthorised use of Generative AI tools are detected, please apply the **DRIVE approach** and refer to the [Faculty Handbook for Academic Violation](#).

Instructors may seek guidance from Associate Deans, and also contact the University Council of Student Conduct for more information:

UNAUTHORISED USE OF GENERATIVE AI TOOLS IN ASSESSMENT

The unauthorised use of generative AI tools will be considered cheating – a violation of the SMU Code of Academic Integrity and will be dealt with accordingly. Please refer to the [Faculty Handbook for Academic Violations](#) for more information (log in required). Penalties will align with the University's established practices for cheating.

Protocol for reviewing suspected cases of unauthorised use of generative AI tools: The DRIVE Approach



The following section outlines the protocol that SMU instructors should follow if they encounter suspected cases of students' unauthorised use of generative AI tools during assessment. The protocol is named the DRIVE approach, which is an acronym for:

D	:	Detect	unauthorised	use
R	:	Review	student	submissions
I	:	Inform	of	status
V	:	Verify	other	sources

E : Escalate via usual academic integrity channels

Details of the DRIVE Approach

49. Artificial Intelligence in Education

The emergence of Generative Artificial Intelligence (AI) based technologies and tools that can generate content including written text, software code and images (e.g., ChatGPT, DALLE2, Midjourney), has generated much interest in higher education while concerns have also been raised about possible misuse, plagiarism, stifling creativity/originality and etc.

SUTD's stance on Generative AI Tools is as follows: the use of Generative AI tools is acceptable when used *ethically and responsibly* to enhance teaching and learning.

SUTD educators have embraced these new technologies and are already applying Generative AI tools in their respective courses to transform teaching and learning in innovative ways. This includes continually reviewing and re-designing class assignments and assessments with the purpose of enabling students to demonstrate capacity for critical thinking, innovation, and originality.

While these developments are taking place, SUTD educators are cognizant of the need to:

- prepare students to use these AI tools ethically and critically; and
- enable students to acquire critical technical knowledge, skills and core competencies needed in their domains.

Here are some guidelines / advisories to help students better navigate in the era of these developments:

1. Students can use Generative AI tools to enhance their learning but not as a replacement for their intellectual contribution.
2. Students can use Generative AI tools in class, assignments, assessments, examinations, etc. only when explicitly permitted by the instructor. Students should adhere strictly to the guidelines provided by the instructor on the usage of such tools.
3. Where permitted by the instructor, students shall disclose and declare how and when Generative AI tools are used in assignments and projects. The format for citation and disclosure of the use of AI tools will be advised by the instructor. Not disclosing or

declaring when Generative AI is used is a breach of academic integrity. When in doubt, students can approach their instructors for guidance.

4. SUTD has clear and established rules on plagiarism, including breaches of academic integrity with the use of AI-assisted tools and apps. Students are bound by the tenets of the SUTD Honour Code* to always conduct themselves with integrity and accountability. The use of AI-assisted tools and apps for all SUTD courses whether in course work or projects (including UROP and Capstone for undergraduate students) to generate content to pass off as one's own work constitutes plagiarism and therefore a breach of academic integrity. Students who are found guilty of plagiarism will be subject to disciplinary action in accordance with the Student Disciplinary Framework.

* Refer to policies on upholding academic integrity at [MyPortal](#) (via > Student Matters and Policies > Student Matters > University Policies > Academic Integrity / Student Discipline).

5. Where students have been expressly informed in their courses or projects that they are accessing confidential information or personal data, students are prohibited from sharing or inputting such confidential information or personal data into Generative AI Tools.
6. To help support students in using Generative AI tools ethically and appropriately, an info-guide titled '[ChatGPT - An Introduction](#)' is available ([view info-guide](#)). The info-guide helps students understand the benefits and limitations of Generative AI tools. Further guides on commonly used Generative AI tools are being prepared and will be released soon.

Concluding

Statement

This is a living document which will be updated alongside developments in the field of Generative AI. Students are encouraged to reach out to instructors for more clarity on the use of generative AI tools in their courses and to be made aware of the benefits, pitfalls and policies governing the use of such AI tools.

50. Generative AI at the Singapore Institute of Technology

Karin AVNIT
SIT Teaching and Learning Academy (STLA),
Singapore Institute of Technology (SIT)

Karin shares her team's experiences of supporting their colleagues in engaging with GenAI, including strategies to enhance professional development and reviewing current assessment practices at SIT.

This post is featured in a Special issue on “Navigating Generative AI in Higher Education”, where academic developers in Singapore’s institutions of higher learning discuss how they are working with their student and faculty colleagues to effectively engage with and navigate GenAI in teaching and learning.

The emergence of generative artificial intelligence (Generative AI, or GenAI) tools like ChatGPT has opened new frontiers in teaching and learning. At the Singapore Institute of Technology (SIT), we identify three focus areas for a responsible adoption of GenAI tools. First, the urgent need to adapt our assessment practices to ensure the validity of our assessment while maintaining standards of academic integrity. Second, the exciting opportunities that GenAI tools provide for improved teaching and learning practices, and third is the need to review how GenAI tools affect the industries for which we prepare our graduates, to update the programme learning outcomes, and to ensure that our students have the right skills as well as knowledge to integrate effectively in the workforce. We share here our experiences and strategies, journeying with academic staff at SIT through this fast-changing landscape for professional development and the review of assessment practices.

A Collaborative Approach to Professional Development

Managing the disruption that GenAI tools introduced starts with ensuring that academic staff are aware of this technology, able to use it effectively, and appreciate its capabilities. This need to quickly upskill all academic staff at SIT is unprecedented in terms of professional development

practices. With the absence of established best practices, no experts that have more than limited experience with the technology in higher education, and with a technology that evolves rapidly, upskilling staff at SIT was a challenge.

SIT took a collaborative approach to this challenge of professional development, turning to the growing collective experience within SIT and drawing insights from the global education community. A channel for discussions on AI in teaching and learning was set up on Microsoft Teams, as a place for sharing information and collaborative learning. In this channel, the team of educational technologists and educational developers from the SIT Teaching and Learning Academy (STLA) published an academic guide for practices related to GenAI. Recognising that any guide on GenAI must be constantly updated, the guide (available only internally) is concurrently published as a shared editable document that enables all teaching and academic staff to contribute from their knowledge and experience. Sharing and dialogue sessions are also scheduled regularly and are well attended.

Guarding the Validity of Assessments

When it comes to the validity of assessments, SIT is perhaps at greater risk than more traditional higher education institutions of GenAI impacting the validity of our assessments. This is due to the progressive assessment practices that we promote and the limited reliance on proctored exams. Authentic and Applied Assessment practices mean that much of the work is done in project or project-like tasks, where instructors have limited oversight or control over the tools used by students in the completion of tasks. For this reason, SIT immediately put an emphasis on reviewing all current assessments to gauge and address any threat to the validity of our assessment practices, and to ensure that we are still able to assess the competencies of students in achieving the stated learning outcomes.

At the same time, we did not change our assessment principles and policies. We still promote authentic assessment practices, rather than react to the potential threat by returning to proctored pen-and-paper exams. In the long run, as GenAI tools also change industry practices, we anticipate that the Programme Learning Outcomes will need to update and reflect the new skills of the

industry for employees to co-work with AI. Then, GenAI tools may be seamlessly integrated into practices of authentic assessment.

As part of the effort to ensure the validity of assessments, all academic staff at SIT are required to demonstrate their competency in evaluating and adjusting assessments to meet the new challenges presented by GenAI tools, as part of a mandatory Micro-Module on AI literacy. As academic staff review and revise their assessment briefs for this demonstration of competency, this naturally feeds back to the collaborative learning to co-exist with GenAI.

Concluding Remarks, as Generated by ChatGPT

SIT's journey with Generative AI in education highlights a path of innovative adaptation and collaborative growth. Emphasising the integrity of assessments, enhancing teaching methods, and aligning with evolving industry standards, we are pioneering a balanced integration of GenAI in higher education. Our experience showcases the potential of AI to enrich learning while maintaining academic excellence, setting an example for the future of educational practices.

51. NUT Position on the Use of Generative Artificial Intelligence in Research

NTU Position on the Use of Generative Artificial Intelligence in Research

Nanyang Technological University (NTU) recognizes that Generative Artificial Intelligence (GAI) – such as OpenAI’s ChatGPT, Microsoft Bing, and Google Bard – provides new tools to support research practices and scholarly activities, including the preparation and drafting of research proposals, manuscripts, and other scholarly work.

NTU acknowledges that GAI has the potential to enhance the quality and efficiency of research, and provide new modes of inquiry. However, it is important to recognize that its use in research can raise ethical, research integrity and other concerns. Researchers must therefore carefully evaluate these issues, and have a thorough understanding of the technology – including its potential

risks and limitations (*e.g., current AI-generated outputs can contain bias, errors, inaccuracies, or falsehoods*) – before considering its use.

Researchers who use AI tools in their scholarly work must use GAI in a responsible and accountable manner, and be transparent on the extent and nature of the involvement of GAI in their work. This would include acknowledging the use of any AI tools in their research proposals, manuscripts, and scholarly works.

Currently, NTU does not impose restrictions on the use of GAI in research, except in the scenarios outlined in points 3 and 4 below.

The Responsible Use of GAI in Research

1. Acknowledging/Declaring the Use of GAI

In the interests of transparency and integrity, the use of GAI beyond basic spelling and grammar checks should be appropriately acknowledged and cited.

Additionally, any use of GAI to generate images for illustrations or figures should be clearly stated in the caption and/or document, in a manner easily understandable to the reader.

2. Authorship

Authorship requires the acceptance of responsibility for the work described in any manuscript. Researchers should recognise that GAI cannot be held responsible as an author for the accuracy, integrity, and content of such work.

Therefore, any GAI (*e.g. ChatGPT*) will not be listed as an author of any paper with an affiliation to NTU; or listed as a Principal Investigator (PI), Co-PI, or collaborator in any research proposals.

Authors and/or PIs are fully responsible for the content of their scholarly materials (*e.g., research proposals, grant applications, manuscripts for publication*) in which GAI was used in the preparation and/or development of such outputs. Researchers must therefore exercise caution and judgement when using GAI, and be ready to verify the accuracy and validity of their work. (*See Figure 1 under resources.*)

3. Data Privacy & Confidentiality

The use of GAI to process or analyse research data must comply with all relevant data privacy and protection laws, regulations, and institutional policies - *e.g., the Personal Data Protection Act (PDPA), NTU's Data Governance Policy.*

Any confidential or sensitive information, and/or personal data are not to be uploaded to any GAI software, system, or platform unless:

- i. Access to the GAI is controlled and restricted to only authorized study members involved in the research;
- ii. The data is not retained in or by the GAI; and
- iii. The activity does not contravene any applicable laws, regulations, or institutional policies in the process.

Researchers must prioritize and safeguard the privacy and confidentiality of research data when using GAI and will be held responsible for any leakage of data.

4. Data Owned by Non-NTU Entities

Data owned by external parties, including but not limited to businesses, organisations, and government ministries and agencies, are also not to be uploaded to any GAI software, system, or platform unless:

- i. All requirements for confidential, sensitive data, and/or personal data listed under “Data Privacy & Confidentiality” are satisfied;
- ii. Valid written permission has been explicitly provided by the data owner/custodian; or

iii. The use of such GAI has been agreed upon in the Research Collaboration Agreement (RCA).

5. Rules established by other Parties

If the proposal, manuscript or other document is being submitted to a agency, journal or other party with rules concerning GAI that differ from the NTU position, then the stricter rules should be followed. NTU researchers may contact RIEO for advice.

In conclusion, Generative AI in research must be used in an ethical, responsible, transparent, and accountable manner that benefits society. NTU faculty, staff, and students should adhere to the following when employing GAI in their research:

- a. Appropriately acknowledge and cite the GAI used;
- b. Do not attribute or list any GAI as an author;
- c. Take full responsibility for the use of any GAI-generated content in their work; and
- d. Safeguard the privacy and confidentiality of confidential and/or sensitive research data, including Personal Data.

Date issued: 4 Jul 2023

Disclosure

OpenAI's ChatGPT (Mar 23 version) was used to improve the clarity and readability of this statement. No confidential or sensitive information was uploaded during this process.

Resources

UNESCO (2023). [ChatGPT and Artificial Intelligence in Higher Education: Quick start guide](#).

52. Guidelines for Collaboration, Co-learning, and Cultivation of AI

Guidelines for Collaboration, Co-learning, and Cultivation of Artificial Intelligence Competencies in University Education1 NTHU AI Task Force National Tsing Hua University Considering the

substantial impact generative artificial intelligence (AI) has on higher education learning environments, we recommend that our institution's faculty and students adopt a transparent and responsible approach when employing AI products, including generative AI, within educational settings: 1 Transparency: Instructors should establish explicit guidelines for AI utilization in their courses, with both students and teachers candidly disclosing their AI usage when relevant. 1 Responsibility: Instructors and students should recognize AI as one among various content sources and possess the ability to evaluate the accuracy of AI-generated content, assuming responsibility for their produced content. We suggest that the role of AI in teaching and learning should encompass collaboration and colearning: 1. Working with AI: Leverage AI tools to collaborate and optimize the benefits of diverse knowledge domains, achieving efficient, innovative, and human welfare-enhancing outcomes. 2. Learning with AI: View AI as a tool for intellectual discovery, supporting the integration and adaptation of various knowledge domains through reflection on learning and thinking processes, enabling the construction of new knowledge forms.

Students and educators should thoroughly understand the following challenges when using AI:

Challenge 1: Technical Barriers: Be aware that language models' complex inner workings hinder the development of open-source, transparent, and democratic models. Challenge 2: Credibility: Identify that scarce training data requires validation of generated content's credibility and accuracy. Challenge 3: Bias and Fairness: Prevent usage of content generated from biased training texts or content that jeopardizes academic integrity. Challenge 4: Pedagogical Methods: Reevaluate teaching and assessment approaches, along with potential outcomes. Challenge 5: Academic Ethics: Examine issues pertaining to authorship and plagiarism. 1 ; English version has been translated by GPT-4 and proofread by Prof. Wing-Kai Hon and Dr. Tonny Menglun Kuo

Integrating AI in Teaching and Learning:

1. Teaching Material Preparation: R AI can support summarization of video subtitles or presentation content for class notes or review. R AI can streamline the creation of presentation materials, such as generating images. R AI can suggest lesson plans (course outlines) and provide diverse, relevant examples as supplementary explanations. 2. Student Learning: R AI can facilitate interdisciplinary learning. R AI can assist in brainstorming, document refinement, and foreign

language editing. R AI can efficiently consolidate key points from literature for focused comprehension and application. R Though generating in-depth reports with AI may be challenging, AI can provide frameworks for students to supplement with their knowledge or perspectives. R AI can serve as a personalized tutor, accommodating individual learning progress and potentially alleviating resource constraints. R Explore "prompt engineering" to leverage AI's integration and retrieval capabilities for innovative idea generation.

3. Educator Teaching:

- R The institution respects educators' AI tool usage strategies in courses; educators should clarify rules for student AI use in course syllabi, especially regarding proper citation and disclosure.
- R Guide students in knowledge provenance, discourage reliance solely on AI-generated content, and address academic and research ethics issues associated with AI tool usage.
- R Emphasize relevant domains' fundamental concepts, promoting deeper learning through extension, association, and application, rather than rote memorization.
- R Investigate AI's potential impact on arts, humanities, and social science disciplines; educators can facilitate group discussions and guide students in delivering oral presentations on AI utilization for creation or proposing alternative perspectives.

4. Learning Assessment (Assignments/Reports):

- R Depending on the course, allow AI assistance for answering questions, but require students to submit assignments in diverse formats, emphasizing domain knowledge construction and knowledge internalization over score attainment.
- R Utilize AI-generated answers as examples, guiding students in critique and revision, annotating edited sentences or paragraphs, and providing justifications.
- R Develop students' ability to trace knowledge origins, assessing sources, accuracy, and school of thought perspectives.
- R Focus grading on students' abilities targeted by the course.

3 5. Learning Assessment (Exams):

- R Design questions requiring deeper reasoning, creativity, analysis (e.g., contextual judgment, pro-con arguments, contentious issues), and critical thinking, emphasizing knowledge's role in problem contexts rather than merely searchable knowledge.
- R Consider AI-human collaboration assessments, permitting AI usage in exams while evaluating students' critical thinking and creativity.
- R For take-home exams, pre-test questions with AI; unsuitable questions may include those on which AI performs well.

Cultivating AI Competence

1. What is AI Competence? AI competence encompasses more than mere programming capabilities; it entails possessing a fundamental comprehension of AI technology and its applications, along with the skills to assess, employ, and execute AI and associated applications. This competence requires the judicious utilization of AI in human learning, work, and life, contemplation of its effects on human existence, and the ultimate promotion of human welfare. Foundational: (1) Understand the basic concepts, techniques, methods, and instrumental nature of AI. (2) Recognize AI application scenarios and potential impacts, and understand the capabilities and limitations of AI tools. (3) Recognize ethical, privacy, and security issues potentially caused by AI and cultivate thinking and judgment abilities. Advanced: (4) Accurately disclose the AI usage process to ensure transparency. (5) Apply various types of AI in work and life situations, improving the ability to solve complex problems. (6) Continuously reflect on the relationship between AI and humans to enhance human welfare.

2. Cultivating AI Competence in Students: Integrating Formal (Courses, Research) and Informal Learning (Clubs, Activities, Competitions, Internships)

(1) Offer foundational and specialized AI courses: Provide (micro) courses to guide students in understanding the basic knowledge, evolution, and risks of AI technologies, helping them comprehend AI's capabilities and limitations, and discern its accuracy; offer specialized AI courses covering topics such as machine learning, natural language processing, robotics, and computer vision to facilitate continuous technical refinement.

(2) Incorporate AI into curricula: Embed AI components within courses across disciplines like engineering, business management, education, arts, and humanities, fostering the development of critical thinking, problem-solving, and innovative collaboration skills. Simultaneously, emphasize the importance of considering social and environmental ramifications in decision-making processes.

(3) Establish academic integrity and accountability mechanisms: Formulate regulations for AI research and applications to ensure compliance with academic ethics and legal requirements.

(4) Provide experiential learning opportunities: Offer students opportunities to participate in AI research projects, construct AI models, and develop AI applications, as well as practical coursework to tackle real-world societal issues.

(5) Forge partnerships with industry collaborators: Enable students to acquire hands-on exposure to real-world AI applications through internships or industry-academic partnerships, thereby understanding the practical implications of technological advancements.

(6) Organize workshops, seminars, or forums to explore AI's impact on various aspects: Provide students with opportunities to learn from experts in different fields about AI's effects on

technology, life, society, academia, and ethics, participate in discussions, and create platforms for peer knowledge sharing. (7) Advocate interdisciplinary collaboration: AI development necessitates the proactive participation of interdisciplinary talents; universities can promote interdisciplinary learning by designing and innovating courses or institutional systems. Additionally, hosting competitions, special projects, or self-study groups can facilitate interdisciplinary and cross-cultural collaboration among students.

53. Guidance for Use of Generative AI Tools for Teaching and Learning

Generative Artificial Intelligence (Generative AI) is a type of artificial intelligence technology that uses machine learning to create new things. Related developments in AI assistance and Human-Robot Collaboration (HRC) are inevitable trends for the future. NTU has adopted a positive attitude towards these trends and encourages faculty members to consider generative AI tools (such as ChatGPT) as an opportunity to enhance teaching. In response to the development of new tools, teachers can make timely adjustments to classroom teaching, and designing content and assessments that better reflect the uniqueness and objectives of your courses. At the same time, students should also understand the limitations of using AI tools, and learn how to use these tools to facilitate your learning.

Faculty and students can gain a better understanding of these tools from the aspects of teaching and learning. Take ChatGPT as an example, we provide the following guidance for the use of generative AI tools.

What is ChatGPT?

ChatGPT is a large language model (LLM) released on November 30, 2022, that is currently receiving a lot of attention. The model is based on generative AI technology. Similar LLMs include PaLM and Bloom. The principle of generative AI is to analyze data models through machine learning, and then automatically generate text, images, audio, video, and other content.

What are the functions of ChatGPT?

The main function of ChatGPT is to perform natural language processing and generate many kinds of textual content, such as:

1. Conversation: conducting a one-on-one dialog between user and AI by answering user's question.

2. Translation: translating text from one language to another.
3. Summarization: condensing long texts into short summaries.
4. Composition: generating articles, stories, news, poems, scripts, etc. based on keywords input by the user.
5. Recommendation: recommending articles, videos, music, etc. based on the user's interests and Internet history.
6. Q&A: answering questions from users, or providing relevant content based on keywords input by users, such as answers to frequently asked questions, knowledge/informational inquiries, etc.
7. Text Analysis: content analysis, topic summaries, character recognition, etc.

What are the limitations of ChatGPT?

1. Network Restriction: It operates on cloud servers and requires an intern connection to use.
2. Language Restriction: Currently, it only supports a limited number of languages and may not be able to process text in certain languages.
3. Data Restriction: Training data is obtained through the Internet which may lead to biases and errors in the generated text, requiring users to make their own judgments and verifications.
4. Privacy Restriction: Some functions may require users to provide personal privacy information, such as chat histories, search histories, etc. Users should pay attention to what data they share.

Guidance for Instructors

How to respond to students using Generative AI Tools in my courses?

If you have concerns about your students using generative AI tools such as ChatGPT, we suggest making three adjustments:

1. **Clearly communicate the methods and restrictions of using generative AI tools to students:** First, you should clarify the principles and guidelines for using generative AI tools in your courses. In addition to verbal explanations and reminders so that students clearly understand the rules, you should also clearly state your rules on generative AI in the course syllabus at the beginning of the course, to help reach consensus with the students and prevent disputes. You should also think about which classroom activities and assignments students are permitted or not permitted to use generative AI tools on. If the

tools can be used, in what forms and scopes are they allowed? If not, how will generative AI use be detected? And how will the misuse be handled?

2. **Enhance classroom practice and exercises:** Depending on the nature of your course, provide students with learning activities that must be performed or completed in the classroom.
3. **Adjust evaluation methods:** **Broaden** the scope of learning assessments. Instead of focusing on a single assessment or end result, look at progress or cumulative results that students demonstrate through the learning process. Another approach is to **deepen** the content of learning assessments by increasing the difficulty of assignments or exams, adding content that is unique to your course, or designing assessments that better reflect students' personal characteristics.

How to use ChatGPT for teaching?

1. Designing exam questions and checking answers: When you create exam questions, try answering them via ChatGPT. If ChatGPT is able to get most of the questions correct, that may mean you need to adjust the questions. Make adjustments by increasing the difficulty of the questions, adding content that is unique to your course, or adding content that reflects students' personal characteristics.
2. Drafting lesson plans or assignment instructions: Try using ChatGPT to create a first draft or to organize your ideas, then refine them, to make your work more efficient. Also use ChatGPT to check whether your assignment instructions are complete and comprehensible, to identify areas that need modification or additional explanation.
3. Creating teaching or learning materials: Use ChatGPT to create worksheets, practice questions, activity instructions, and resources for student self-learning.

What courses and learning content are suitable for using ChatGPT?

ChatGPT is a language model that can be used to support a variety of courses that require natural language processing. Examples include:

1. Language and Linguistics: Help students learn and analyze various aspects of language, including grammar, syntax, and semantics.
2. Communication and Media Studies: Explore how language is used in different forms of media and communication channels, including social media, news, and advertising.

3. Computer Science and Artificial Intelligence: Help students understand the principles of natural language processing and machine learning, including techniques for text classification, language modeling, and sentiment analysis.
4. Psychology and Cognitive Science: Study how humans process language and communicate with each other, and how technology is changing communication patterns.
5. Education and Language Studies: Support language learning by providing feedback, practice, and dialogue.

Are there any tools to detect whether or not a student is using ChatGPT?

Currently, detection tools for ChatGPT-generated content include OpenAI AI Text Classifier, CheckforAI, GPTzero.me, and Content at Scale. However, we need to remind you that the current technology for detecting ChatGPT-generated text does not provide sufficient accuracy for you to say for certain whether AI-generated content has been used in a particular assignment. In addition, detection is more difficult if the generated content has been modified rather than used directly. Furthermore, since the content of AI-generated text is arbitrary combinations of words, even if a detection tool determines that AI-generated text was used in an assignment, it cannot provide conclusive evidence (that is, unlike existing plagiarism detection tools that can clearly identify the journals or web pages containing similar content). This being the case, teachers must be cautious about using any AI content detection tools, to avoid false detections that cannot be validated and that could lead to further disputes.

Will ChatGPT replace introductory-level courses?

The existence of ChatGPT does not diminish the value of learning the fundamentals of any given topic. Responses currently given by ChatGPT often have errors and do not give more in-depth reasoning or creative production. The majority of research indicates that while the quality of ChatGPT-generated content is moderate to high, it is not yet comparable to leading-edge results. When students (or anyone) work with the content provided by ChatGPT, they must first be able to identify the correctness of the content, evaluate the quality of the results, and know how to apply the relevant content. Additionally, using ChatGPT requires that the user describe their needs in a correct, logical manner. All of this requires a high level of training and knowledge in the fundamentals.

Guidance for Students

How to use content generated by ChatGPT?

When you use ChatGPT to facilitate writing assignments or reports, you must clearly indicate that content produced by ChatGPT is used, so that the reader understands what resources you have used to support your argument. If you find that the content used comes from the works of others, use an appropriate citation format, such as APA, MLA, Chicago Manual of Style, etc. (as required by your School or course) to indicate the exact source of the content. When you use AI-generated content, it is necessary to conduct information verification; ensure that academic ethics and academic integrity requirements are followed, and that you don't engage in plagiarism or copyright violations.

How to cite content generated by ChatGPT?

At present, there are no definite rules for citing AI-generated content in academic writing. However, the sources of AI-generated content cannot be retrieved, accessed, or provided with direct links. Therefore, we recommend treating it as “personal communication” or “correspondence” and make use of the appropriate citation format. For details and citation methods, please refer to APA, MLA, Chicago Manual of Style, and other citation format. You can also refer to Scribbr for citation suggestions and examples: [ChatGPT Citations | Formats & Examples](#) .

How to use ChatGPT to help with learning?

1. Answering questions: Ask ChatGPT questions to further clarify concepts or ideas that you're unclear about within a topic of study or course content.
2. Providing feedback: Use ChatGPT to analyze your assignments and provide feedback as a reference for reflection and improvement.
3. Providing Examples: Use ChatGPT generated content as examples to compare or analyze the strengths and weaknesses of your own work.
4. Editing written content: Use ChatGPT to assist with coursework or reports writing by correcting grammar or refining content.
5. Practicing languages: Use the conversation and question-and-answer functions for learning foreign languages.
6. Practicing critical thinking: Content generated by ChatGPT may include information bias and errors. It is thus necessary to learn to examine, evaluate, and analyze information. With content that lacks supporting evidence or seems incorrect, seek other authoritative sources

to ensure that the information is accurate, reliable, and meets academic and research requirements

7. Clarifying ideas and identifying goals: When you use ChatGPT, your questions must be as clear and specific as possible to receive high-quality responses. The question-and-answer process with ChatGPT can thus help you clarify your ideas and better identify your learning goals.

What should be aware of when using ChatGPT?

1. Avoid over-reliance on ChatGPT: The content generated by ChatGPT can only be used as a reference; it cannot replace personal thinking and research results. Therefore, you cannot rely on ChatGPT-generated content for learning, or even research. Instead, use the content generated by ChatGPT as a reference and combine it with personal analysis and reflection to complete your learning.
2. Improving the quality of questions: To get quality answers from AI-generated content, you must first have quality questions. The questions you ask must be as clear and specific as possible, so that ChatGPT can interpret them more readily and provide meaningful answers. Even then, ChatGPT may still produce incorrect or ambiguous answers, so it is up to you to assess the correctness of the answers.
3. Clearly defining your learning objectives: It is important to have clear learning objectives before using ChatGPT. That way, you can focus on questions and answers related to these objectives, and deepen your understanding of the topics through the information you obtain from ChatGPT. In addition to checking the information yourself, we suggest that you discuss the information with other students and teachers. By obtaining different perspectives and opinions, you can confirm that the information you're using is consistent with your learning objectives.

54. Use of Artificial Intelligence Tools

THE CHINESE UNIVERSITY OF HONG KONG Use of Artificial Intelligence Tools in Teaching, Learning and Assessments A Guide for Students Artificial intelligence (AI) is sweeping the globe, and generative AI in particular has been hotly discussed because of its potential to

revolutionize the way we teach and learn. The University believes that it is crucial for teachers and students to embrace and become acquainted with AI in order to optimize its potential in education. Students should learn to make sensible use of AI tools, not only for their studies, but also for their future professional development and advancement in order to thrive in this AI era. Like any other educational resources, teachers and students should approach AI tools critically, recognising their limitations in an honest and authentic manner and how these tools could be incorporated into teaching and learning in order to attain the desired learning outcomes. To optimize the use of AI in education, the University has prepared for teachers the “Guidelines on the Use of Artificial Intelligence Tools in Teaching, Learning and Assessments” (Guidelines) to i) set out how the University may integrate AI tools in its teaching and learning while upholding academic honesty, integrity and quality; ii) recommend some possible approaches in adopting AI tools in teaching and learning; iii) make it clear and explicit that improper/unauthorized use of AI tools in assignments/assessments constitute acts of academic dishonesty which will be handled in accordance with the University’s existing guidelines and procedures; and iv) provide some readily available references and resources for supporting the adoption of AI tools in teaching and learning. Students should take note of the following salient points extracted from the Guidelines and follow strictly the instruction and/or permission given in the course outline by the teachers regarding the use of AI tools in teaching, learning and assessments. The Guidelines will be reviewed and updated as needed to reflect changes in technology, best practices, and other relevant developments.

a) There are different types of AI tools, for instance generative AI tools (e.g. Chat GPT) which can be easily instructed using ordinary human language to generate various formats of texts. Some AI tools facilitate the creation of ‘original’ artwork (e.g. DALL·E 2), translated text (e.g. Google Translate), formulas (e.g. Sheet+), and computer code (e.g. OpenAI Codex), etc. applicable to a great variety of use. While teachers and students are encouraged to explore and take advantage of the benefits of adopting appropriate AI tools to enhance their teaching and learning activities, decisions on which AI tools to adopt and how to use them in teaching and learning should be made cautiously and thoroughly.

b) The availability and accessibility of AI tools to students will be carefully evaluated before adopting AI tools into any teaching and learning activities. AI is a double-edged sword; we should use but not abuse it, use it as a research but not cheating tool, and most importantly, use AI to think with you, but not for you.

c) Where applicable and permitted, approaches to the use of AI tools in different disciplines will be worked out taking into

consideration the needs of different disciplines, their pedagogical approaches and assessment means. When adopting AI tools in teaching and learning, teachers and students should be cautious of their accuracy and reliability and bear the responsibility of using the educational resources and references obtained through these tools.

2 d) As a general principle, students are prohibited from using any AI tools to complete their assignments, assessments and any other works that count towards their final grade of the course or attainment of the desired learning outcomes, unless explicitly permitted.

e) Depending on the learning outcomes, pedagogical design, and assessment scheme of different courses, the following are some possible approaches to adopt AI tools in teaching and learning. Relevant details will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction and permission strictly and seek clarification from the course teacher if in doubt. Students are also expected to understand the limits and appropriate uses of these tools.

i. Approach 1 (by default) - Prohibit all use of AI tools

In assessing the level of achievement of learning outcomes and students' performance, students are expected to produce their own work independently without any collaboration or the use of AI tools. That says students are prohibited from using any AI tools in their assignments and assessments that count towards students' final grade of the course, or for evaluating their attainment of the desired learning outcomes.

ii. Approach 2 - Use only with prior permission

In some courses, it may be appropriate to use AI tools in some in-class exercises or assignments. Where applicable and permitted, students will be clearly and explicitly informed of when and how they can use these tools which shall be cited or acknowledged in their work. Details will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction and permission strictly and are expected to understand the limits and appropriate uses of these tools.

iii. Approach 3 - Use only with explicit acknowledgement

In courses where students are allowed or expected to collaborate with or use AI tools, students may use these tools for in-class learning activities, exercises or assignments as long as they explicitly cite or acknowledge the use of these tools. Details will be spelt out clearly in the course outline and/or the instructions of the assignments. Students shall follow the instruction strictly and are expected to understand the limits and appropriate uses of these tools.

iv. Approach 4 - Use is freely permitted with no acknowledgement

In courses where students are allowed or expected to frequently collaborate with or use AI tools, students may use these tools for in-class learning activities, exercises or assignments without citing or acknowledging the use of these tools. In these classes,

it is critical that students understand the limits and appropriate uses of these tools. Details on which AI tools are to be used will be spelt out clearly in the course outline and/or the instruction of the assignments. Students shall follow the instruction strictly and are expected to understand the limits and appropriate uses of these tools. f) The adoption of permitted use of AI tools in courses are subject to regular review by the course teacher(s) and the programme concerned. It is the responsibility of students to study the course outline, assessment scheme and instruction of individual assignments in detail to ensure that they follow the instruction and permission strictly. Improper/unauthorized use of AI tools in learning 3 activities and assessments constitute acts of academic dishonesty which will be handled in accordance with the University's Procedures for Handling Cases of Academic Dishonesty. g) As a general principle, students are expected to complete assignments/assessments on their own without any external assistance, unless otherwise specified. If AI tools are permitted for use in the course, students should pay attention to the following for proper use of these tools: i. students should learn and use these tools responsibly and ethically, and be aware of their limitations; ii. the quality of output of some generative AI tools correlates directly to the quality of input, students should master “prompt engineering” by refining their prompts in order to get good outcomes; iii. students should fact-check all outputs of AI tools by cross-checking the claims with reliable sources and are responsible for any errors or omissions, if any, when using these tools; iv. like any other tools and references, permitted use of AI tools should be acknowledged unless otherwise specified; specific and detailed information on the AI tools used, including prompts used if applicable, for completing the assignments/assessments should be provided in the work concerned and, if deemed necessary, the output of generative AI should be included as an appendix of the work submitted by students. h) As a general principle, students are expected to complete all coursework, formative and summative assessments independently without the use of AI tools or other forms of unauthorized assistance, unless specifically permitted. Improper and unauthorized use of AI tools not only jeopardize the quality and efficacy of teaching and learning, but they also constitute acts of academic dishonesty. Students should be cautious of the following which may result in improper/unauthorized use of AI tools in learning: i. using AI tools in completing assignments/assessments without prior permission; ii. handing in an AI-generated work as one's own; iii. using AI tools to cheat in a course; iv. using AI tools that are not up to date and result in the use of outdated and inaccurate resources; and v. using AI tools in an unethical and irresponsible manner. i) If the use of AI tools

is not permitted in an assignment/assessment, and a student is later found to have used such a tool in the assignment/assessment, the case should be handled in accordance with the University's Procedures for Handling Cases of Academic Dishonesty. j) Similar to the submission of other assignments, students are required to declare and assure that the works submitted are their original works except for source material explicitly acknowledged, and the permitted use of AI tools in the assignment(s), if applicable. The academic honesty declaration statement is updated accordingly in the VeriGuide System. k) While appropriate enhancement would be implemented to the VeriGuide system for detecting cases of improper/unauthorized use of AI tools, the University will explore and devise other appropriate measures in detecting students' submissions and assessments. All identified cases will be handled in accordance with the University's Procedures for Handling Cases of Academic Dishonesty. l) Similar to other serious cases of academic dishonesty, penalties for improper/unauthorized use of AI tools in assignments/assessments may include reviewable/permanent demerit(s), failure grade 4 for the course concerned, suspension from the University, lowering the degree classification, and termination of studies at the University. m) For cases in which improper/unauthorized use of AI tools in assignments/assessments was committed by a person when s/he was still a student of the University studying for the award in question (but only came to light after graduation) may also constitute good cause under which the University may revoke his/her academic award in accordance with the University's procedure for revoking an academic award. n) The educational needs of students in the use of AI tools will be reviewed. o) University will provide professional development opportunities and support to teachers to enhance their AI literacy so as to help enhance sensible and critical use of AI tools among students. p) Guidelines and policies on the use and application of AI tools in teaching and learning will be regularly reviewed and updated to reflect changes in technology, best practices, and other relevant developments.

55. Chulalongkorn University Principles and Guidelines for using AI Tools

Principles

1. Chulalongkorn University supports its students and staff in gaining understanding and being able to appropriately use AI tools.

2. Chulalongkorn University supports the modification of teaching and assessment processes, as well as work tasks, appropriate for the creative and ethical use of AI tools.
3. Chulalongkorn University places the utmost emphasis on academic integrity.

Guidelines

1. Teaching and Assessment
 - Instructors should understand the capabilities and limitations of the chosen AI tools, and design teaching and assessment processes suitable for these tools.
 - Instructors should clearly define in the course syllabus the scope and directions for using AI tools, such as ChatGPT, Google Bard, in the course, along with explaining the reasons to the students.
 - If a course allows students to use AI tools, the instructor should adjust the evaluation methods appropriately, and should not directly assess the work that students can use AI tools to respond to.
2. Usage of AI Tools
 - AI is merely a tool. The user is responsible for the correctness of the information and resulting work.
 - If AI tools are used in any task, reference and clearly state the scope of usage in that task. Concealing and not disclosing information about using AI tools is considered an ethical violation, which could be penalized according to relevant regulations and rules.
3. Privacy and Personal Information
 - It is prohibited that any confidential information should be uploaded from the organization or personal information into the AI system without consent from the concerned individuals.

56. ChatGPT General Usage

The use of AI in academia presents both opportunities and challenges to academic integrity. Concerns include the use of AI-generated papers and the potential for biased or inaccurate analysis. However, guidelines and protocols can be developed to ensure ethical use of AI in academic work.

It is important for academic institutions, researchers, and students to work together to maintain academic integrity while utilizing the transformative power of AI in education and research. Lecturers, do you... 1. have to ask students to STOP students using ChatGPT? NO. As a teacher and a researcher, we should embrace the use of technology. However, as highlighted by Mr Nadella, a SAFETY element must be put in place. For example, lecturers should inform the students about academic integrity and highlight the rules and regulation in UM (refer to Kaedah-Kaedah Universiti Malaya (Tata tertib Pelajar) 1999) if students abuse the technology. BANNING ChatGPT is not the solution! 2. know that there are other AI that can help you to counter ChatGPT? You may use ChatGPT counter-AI to help you identify 'cheating' students. Or you can simply use alternative mode of assessment for your course. 3. provide resources and support to help students complete their coursework without resorting to plagiarism? Students may use ChatGPT to complete their assignment due to lack of understanding, motivation or even because they desire to get good grades easily without putting in the necessary effort. As such you may also provide necessary support by providing additional learning materials, or perhaps provide "sentences" that student may use to ask for info from ChatGPT. 4. use plagiarism software to check for students work? As of this writing, many plagiarism tool DO NOT detect AI generated text, including the most famous – TurnItIn and Grammarly. Lecturers should be mindful when giving written-based assignment to students. world, understanding, learning human preferences, with safety..."

57. Guide for ChatGPT usage in Teaching and Learning

ChatGPT is a natural language processing (NLP) tool that uses a type of AI called a GPT-3 model to generate responses to text-based chat input. It is designed to simulate human conversation and provide responses that are relevant to the user's input. It can be used for a variety of applications, including customer service, personal assistants, and chatbots.

As an educator, it is important to have a general understanding of artificial intelligence and how it can be used in education. Students will have to learn how to navigate life with AI. Usage of AI in teaching and learning should be conducted in a safe and ethical manner, and with integrity. ChatGPT usage requires educators to redesign their assessment. This technology will eventually start to push some of those terrible assignments out and force us to come up with something new LET STUDENT USE GPT Allow students to use ChatGPT and hold discussions according to established rules.

ENSURE STUDENTS UNDERSTAND Practice recall and other memorization exercises that mention a specific period of time, a particular subject, or past actions are used to make sure that students make an attempt to comprehend and evaluate any references they have used.

COLLABORATIVE ACTIVITIES Increase the number of discussions and collaborative activities. Students use their own short- and long-term memories when they converse. While individuals can search up short answers, most of the effort involved in having a discussion comes from their own thoughts. Students can summarise a discussion and share their comments on it afterward, which is much more difficult to do with a bot.

ENGAGE WITH STUDENT Put an emphasis on experiential learning and include students in personalized elaborations that are related to their daily routines and immediate surroundings. Allow students to demonstrate their learning. It will be more challenging (but not impossible) for students to simply ask an AI to write their assignment if you require them to incorporate ideas, evidence, viewpoints, and data from current or personal events or geographical contexts.

UNIQUE CASE STUDY Use AI to your advantage. For your authentic evaluation, utilise ChatGPT to tailor or create a special case study for each student or their group based on their interests and proficiency level.

DIFFERENT TYPE OF ASSESSMENT Try utilising various assessment modalities, such as oral evaluation and live demonstration, that are more resistant to AI and can help students grow and exhibit comprehension. Furthermore, generative AI poses less of a threat to staging exams, which require students to submit draft, get feedback, and then revise their work. Create questions that involve :- contextual and real world problem (eg relate to a specific issue/discussion/topic in a past conducted lesson) personal opinion with reasoning and justification current issues drafting sketches/figures/illustrations

STUDENT OPINIONS AND JUSTIFICATIONS Include a section in assessments to allow students to give feedback on how they can improve (what they got, if it fits, how to organise it, how to communicate it effectively, etc.) and reflect their synthesis of the information they have learned through reading, internet research, peer discussions, and ChatGPT responses they have used.

58. Guidelines for Responsible Research Conduct (2023)

Table of Contents	1. Research Topic Selection and Implementation.....
.....	12. Data Management

.....33. Literature citations.....64. Achievements
signature.....75. Results published.....	
.....86. Peer review.....	
.....12 7. Ethical review.....148.
Academic exchanges and cooperation..... 179.
Intellectual property protection..... 1810. Training and
guidance..... 1911. Supervision and
management..... twenty one	

Guidelines for Responsible Research Conduct (2023) Responsible scientific research is an important guarantee for promoting the healthy development of science and technology, and is an inevitable requirement for achieving high-level scientific and technological self-reliance. In order to guide scientific researchers and scientific research institutions, universities, medical and health institutions, enterprises, etc. (hereinafter collectively referred to as "scientific research institutions") to carry out responsible scientific research, the Supervision Department of the Ministry of Science and Technology organized the preparation of the "Guidelines for Responsible Research Conduct" and proposed Scientific ethics and academic research norms that should be generally followed in scientific research practice.

1. Research Topic Selection and Implementation

Research topics should adhere to the "four aspects", highlight the problem orientation, comply with the requirements of science and technology ethics and science and technology safety regulations, avoid simple duplication or low-level research, avoid being divorced from reality or blindly pursuing hot topics, and must not conduct research prohibited by laws and regulations. The application, implementation and final acceptance of research projects should comply with relevant regulations, and necessary resources should be invested in the research process.

(1) Scientific researchers

1. The selection of research topics should consider the scientificity, innovation, applicability and feasibility of the research. After sufficient literature review, investigation and scientific demonstration, combined with the necessary resource conditions to complete the research, the

difficulties and doubts in the research field should be sought. , blank spot. If the purpose is not verification, you should avoid simply repeating research that has been carried out by others.

2. Application materials for research projects should be true, accurate and objective. The same or similar research content shall not be used for repeated declarations, and others shall not be listed as members of the research team without consent. It is not allowed to plagiarize, buy, sell or ghost-write application materials, and it is not allowed to use generative artificial intelligence to directly generate application materials.

3. Research funding budgets should be prepared in accordance with relevant regulations and actual needs, and research funding should be used rationally and prudently. Expenses for completed research or expenses unrelated to the applied project must not be included in the budget.

4. It is not allowed to seek improper benefits from the review organizer, its staff, review experts, etc., directly or indirectly, explicitly or implicitly. 5. The time and energy required for research should be ensured, and the number of projects hosted and mainly participated in during the same period should comply with the restrictions of the relevant funders. The implementation of research should be rigorous and scientific, with appropriate methods and technical means. The mission statement or contract obligations should be strictly performed. The target tasks and agreed requirements should not be arbitrarily reduced. Research tasks should not be subcontracted or subcontracted to others in violation of regulations, and irrelevant research results should not be used as offsets. Crossover. 6. In activities such as research topic selection, project application, research implementation, project acceptance, and participation in project review as an expert, possible conflicts of interest should be promptly identified and proactively declared to avoid affecting the objectivity and impartiality of the research or review, and damaging the interests of others or society.

7. Relevant regulations on security and confidentiality, fund use, resource and data sharing, and intellectual property ownership should be strictly observed. Use generative artificial intelligence appropriately and in compliance with regulations to participate in research implementation.

(2) Scientific research units

1. Review and check the scientific research integrity of the project applicant and the authenticity and completeness of the application materials, and make good scientific research integrity a prerequisite for application.

2. Fulfill the main responsibilities, provide necessary supporting conditions, track research progress in a timely manner, and strengthen full-process management of research activities.

2. Data management

The collection and recording of research data should be complete, accurate, and traceable, and the storage and use should comply with professional norms and management regulations to ensure data quality and data security.

(1) Scientific researchers

1. Data collection operating specifications should be strictly implemented, appropriate data collection methods should be selected, and research data should be collected objectively, completely, and accurately. 2. Record the research process and experimental data promptly and accurately to ensure the integrity, objectivity and traceability of research records and data. Data must not be recorded or used selectively to obtain specific results.

3. If written records are used, an experimental record book with continuous page numbers and that meets the storage requirements should be selected; the original data, charts, photos, etc. generated by the experiment should be pasted in the corresponding positions of the experimental record book in an orderly manner and marked in detail. Corrections to the record should be made by the original recorder, and the original content must not be obscured. The reason for the correction should be stated and signed. Do not alter data or destroy any part of the notebook. It is not allowed to fabricate or tamper with original data, or save artificially processed data as original data. If electronic records are used, they should be combined with experimental records

Correlate and ensure that the data and generation time have not been artificially changed. 4. If you use data that has not been officially published by others, you must obtain the consent of the data owner in advance and explain the source of the data.

5. The use, dissemination, copying, storage, deletion, etc. of data should comply with the requirements of the "Data Security Law", "Personal Information Protection Law", "Scientific Data Management Measures", etc. The collection of human experiment data or sensitive data involving privacy and confidentiality should strictly abide by relevant laws, regulations and ethical norms, and can only be carried out after obtaining the informed consent of the relevant personnel or the approval of an agency with approval authority. The data may not be used for purposes other than

the agreed purpose, or the data may be transferred or disclosed to other institutions or persons without consent. 6. When analyzing and processing data, appropriate analysis methods and processing methods should be used to comprehensively, clearly and accurately reflect the research process and research results, and explain them in detail in the research report.

7. When processing academic images, you should comply with the specifications of the corresponding discipline or academic publishing unit, and indicate the portion of the academic image that is processed when publishing. The key information contained in the original academic image shall not be changed, blurred, eliminated or distorted, and the academic image shall not be subjected to inappropriate or deceptive processing, including adding, removing or moving objects, removing or blurring the background, etc. The spliced combined pictures should add obvious dividing lines between each group of pictures, or clearly explain the splicing of the pictures. Other pictures not obtained under the conditions of this experiment may not be used to replace the real experimental pictures, and the source or source must be indicated to directly use pictures from other people's research results.

8. After the research results are published, it is recommended that they be published without violating confidentiality and intellectual property regulations.

Under this circumstance, the original data involved, as well as methods, reagents, software, source code and other materials involved should be submitted or openly shared in an appropriate manner to improve the application value of the data. 9. Organize, save, and back up research data in a timely manner, and take effective measures to prevent data from being lost, leaked, damaged, or tampered with.

10. Properly preserve all experimental records, experimental data (including unpublished data, negative data, etc.) and experimental notebooks in accordance with the regulations of the subject field or the scientific research unit where you are located, and preserve experimental samples in accordance with relevant technical specifications. Within one month after the publication of research results such as papers, the experimental records, experimental data and other original materials involved should be submitted to the scientific research unit for centralized archiving, or in accordance with the relevant management regulations of the scientific research unit. 11. The project leader should supervise the collection, storage and utilization of data, conduct necessary inspections or verifications on the collected data to ensure that the data is reliable, and properly preserve all records and original data within the specified period. 12. When it is discovered from

the research data that there may be a serious impact or threat to public health, ecological environment, public security or social order, it should be promptly and proactively reported to the relevant departments. Data release should comply with relevant regulations, remain transparent and objective, and avoid intentionally highlighting, emphasizing or concealing or ignoring specific content. 13. Relevant laws, regulations and academic norms should be followed, and generative artificial intelligence should be used reasonably to process text, data or academic images in accordance with regulations to prevent risks such as forgery and tampering with data. 14. It is not allowed to obtain experimental research data by paying third-party organizations or others without actually conducting the research. When the conditions are not met to entrust a third-party organization to conduct experiments or data collection, the scientific researchers should propose an experimental design plan and Conduct analysis and research based on original experimental records and data provided by third-party institutions, and the source of the data should be stated when publishing. The use of third-party survey statistics or relevant public database data should be obtained through legal channels, and the source or source should be indicated. (2) Scientific research units

1. Establish a management system and quality control system for research data, and clearly stipulate and strictly implement the collection, collection, storage, ownership, use, sharing, confidentiality, and security of data. Regular inspections are carried out to ensure that the original records of research activities are timely, accurate, complete, properly preserved, and queryable and traceable. 2. Provide guarantees in terms of software and hardware facilities, funds and personnel required for research data storage, management, services and security. According to the characteristics of research activities, original recording media in a unified format, continuously numbered, and easy to use are produced and properly preserved and archived. 3. Literature citations

Researchers should clearly distinguish their own research results from those of others in project applications, final reports, papers or other research results. When referring to other people's opinions or research results, they should be realistic and accurate, and indicate the source in an appropriate manner. 1. If you refer to or draw on other people's academic opinions, research ideas, or published works in your research, you should indicate it in the form of citations, notes, acknowledgments, etc. in accordance with the general standards or norms of this discipline.

2. Quote original documents as much as possible. If it is necessary to cite original documents cited or summarized by other authors due to inability to obtain original documents or other reasons, they should be marked as quotations and strive to be accurate.
3. Be objective and accurate when quoting other people's documents, and avoid misquoting or quoting out of context. You should not cite results in unfamiliar research fields, you should not cite without understanding the research content or progress, and you should not intentionally distort, elevate, or belittle other people's academic opinions or research findings. In principle, retracted articles will not be cited except for critical purposes.
4. Content generated using generative artificial intelligence, especially when it involves key content such as facts and opinions, should be clearly labeled and explained its generation process to ensure authenticity and accuracy and respect for the intellectual property rights of others. Content that has been marked as generated by artificial intelligence by other authors should generally not be cited as original literature. If it is indeed necessary to cite it, an explanation should be given.
5. If you use charts or pictures that have been published by others, you should obtain permission from the copyright owner in advance and use it within the scope of permission, and indicate the source or source.
6. You should not deliberately ignore or conceal relevant and important literature published by others or literature that is detrimental to your own research.
7. Documents that have not been referenced or are not relevant to the research content should not be included in the references, including inappropriate self-citation, cross-citation in agreement with others, or citation of irrelevant documents according to the requirements of reviewers and editors, etc. References generated by generative AI may not be used directly without verification.

4. Signature of results

The signer should have made substantial academic contributions, that is, made important contributions to research ideas, design, implementation, data acquisition, data analysis and interpretation, etc., or made critical revisions to important intellectual content. Those who have no substantial academic contribution to the results shall not be signed.

1. All signatories should review and agree in advance to publish the results with their signatures, and are responsible for the content that they have completed or participated in. The first completer, first author and corresponding author bear primary responsibility for the authenticity and reliability of the results.
2. All signatories should make substantial academic contributions to the results and may not be named or impersonated. If you find that you have been named or impersonated, you

should take the initiative to raise questions and request corrections. 3. The order of signatures is usually arranged according to the size of the contribution to the results. It should generally be decided jointly by all those who completed the results, or follow the signature conventions of relevant disciplines. 4. For individuals and organizations that are not eligible for authorship but have contributed or helped the research work, their contributions can be explained through acknowledgments, notes, etc. 5. For funded research results, the funding agency, project name and approval number should be truthfully labeled as required. For project results funded by multiple institutions, in principle, they will be sorted by the size of their contribution to the results. Projects that are irrelevant to research work or that are false shall not be marked. Do not withhold funding sources, withhold true authorship information, or fictitious authorship to conceal a conflict of interest.

6. The main unit that completes the results shall not be changed at will due to changes in the work unit of the person who completed the results. Personal information such as work unit and job title shall not be fabricated or falsified. 7. Generative artificial intelligence shall not be listed as a co-contributor of the results. The main methods and details of using generative AI should be disclosed in relevant places such as research methods or appendices. 5. Publication of results

When publishing or distributing papers, monographs and other results, you should fully state the research process, clearly introduce the research methods, accurately describe the research conclusions, and submit or share relevant data as required to facilitate others to repeatedly verify and judge the reliability of the research results. (1) Scientific researchers

1. Research results should first be published through the peer review process, or communicated within the scientific community through academic reports, academic seminars, preprints, etc. Publication of breakthrough research results and major research progress must be approved by the scientific research institution where the applicant is located. Research results that have not been scientifically verified or peer reviewed may not be disseminated to the public. 2. The publishing unit of the results to be published, the database containing the results, etc. should be verified to avoid publishing units that lack quality assurance or are false. 3. The same manuscript reporting research results or manuscripts based on the same data with only minor differences may not be submitted to two or more publishing units for publication at the same time. Only after receiving a rejection notice, or exceeding the prescribed review period, or applying to withdraw the submission and receiving confirmation from the publishing unit, can the manuscript be transferred

to another publishing unit. If the work is jointly completed by multiple authors, the consent of all authors must be obtained before the decision is made to resubmit. 4. Published papers or their data, pictures, etc. are not allowed to be republished, and parts of multiple published papers are not allowed to be pieced together to create "new results" for publication. If it is really necessary to re-publish, permission must be obtained from the published and to-be-published publishing units in advance. When re-publishing, an explanation should be made in a prominent position and the original publication location should be noted. 5. A complete research result shall not be split into several results for publication in order to maintain the integrity, systematicness, scientificity and logic of the relevant results, and results signed by different authors based on the same research content shall not be written. 6. The first completer, first author and corresponding author are responsible for ensuring that all signatories agree to publish and recognize the final results.

7. When publishing results, you should declare conflicts of interest and indicate the source of funding for the research results as required.

(2) Scientific research units

1. Strengthen the management of research results published by scientific researchers, and establish systems such as integrity commitments for publishing research results, collection and verification of original data, traceability of research processes, inspection and reporting of research results, etc. For projects supported by fiscal funds, scientific researchers should be urged to submit relevant scientific and technological reports as required. For those who publish multiple papers, obtain multiple patents and other achievements in a short period of time, verification and verification should be strengthened in accordance with regulations. Results that scientific researchers plan to publish involving confidential or sensitive information will be reviewed and checked in accordance with relevant regulations. 2. Those who publish research results in academic journals that are included in the early warning list should be promptly warned; research results published in academic journals that are included in the "blacklist" will not be recognized in various reviews and evaluations. , and related publication expenses will not be reimbursed. 3. If the academic papers and other research results published by the scientific researchers of the unit violate the requirements of scientific research integrity, the relevant responsible persons should be severely dealt with and required to take measures such as withdrawing the papers.

(3) Academic publishing units

1. Establish and improve systems such as peer review, ethics supervision, copyright management, academic standard commitment, handling of objections to retractions, and conflict of interest management, and establish a management and supervision system for editors, editorial board members, and reviewers. 2. Clarify the publication standards of research results through the "Call for Papers" and "Instructions for Authors", etc., and may require the contribution of each signatory to be stated when publishing the results. 3. Authors should be required to disclose whether they use generative artificial intelligence, indicate the specific software name, version and time of use, and make specific annotations for auxiliary generated content involving citations of facts and opinions.
4. Records of original submissions, review comments, revised drafts, correspondence, and editorial rejection or acceptance decisions of published research results should be kept for at least 3 years for verification. 5. Scientific research dishonesty in authors' submissions should be detected and screened, and complaints and reports about scientific research dishonesty should be accepted. For manuscripts with scientific research dishonesty or serious errors, measures such as commenting, attention, correction, and withdrawal should be taken, and the relevant databases that include the research results and the author's unit should be promptly informed. 6. Reasonably select reviewers, urge reviewers to conduct serious and fair reviews, and conduct corresponding supervision and evaluation of their compliance with scientific research integrity requirements. Reviewers are reminded to be cautious when using generative artificial intelligence during the review process. You shall not exert undue influence on the reviewer's compliance review, or negate or distort the reviewer's review opinion without reason. 7. Editors and reviewers are not allowed to disclose, publicly discuss, plagiarize or appropriate the author's unpublished research results without authorization; they are not allowed to deliberately delay the review or publication progress; they are not allowed to take advantage of the publishing process and the content of the manuscript to seek improper interests; and they are not allowed to increase the influence of the journal. Authors are asked to cite specific literature where necessary.
8. Comply with relevant regulatory requirements on conflicts of interest. Editors should be required not to conceal their interest relationships with contributors, or intentionally select reviewers with potential or actual interests or conflicts of interest to review manuscripts. 6. Peer review

Peer review is an important reference for research resource allocation, acceptance of research results, publication of research results, talent evaluation, science and technology awards, professional title promotion, etc. Review activity organizers and reviewers should ensure the scientificity, authority, objectivity and fairness of the peer review process and create a good peer review ecosystem. Researchers should actively participate in peer review activities.

3. Reviewer

1. Carry out review work objectively, fairly and rigorously, respect the dignity and academic autonomy of the reviewees, respect different academic viewpoints, and put forward review opinions and suggestions constructively. Comments must not contain words or comments that are insulting, deliberately disparaging, or unfair. Comments should not be influenced by non-academic factors.
2. Participants should participate in peer reviews based on their own professional knowledge and abilities, and should not participate in review work that is unfamiliar to the matters being reviewed or related research directions or that cannot be completed within the specified time.
3. Provide specific and detailed comments and explain reasons or evidence when necessary. You are not allowed to ask others to review or write reviews on your behalf.
4. If there is a conflict of interest with the person being reviewed, you should take the initiative to explain the situation to the review organizer, and actively recuse yourself as required or let the review organizer decide whether to participate in the relevant review.
5. Strictly abide by the disciplines of the review work, and are not allowed to accept requests such as "saying hello" or "building relationships", and are not allowed to ask for or accept gifts or other gifts from stakeholders.
6. Keep the review content and process confidential as required. Do not copy or spread the reviewed materials without authorization, and do not disclose information such as review objects, expert opinions, review conclusions, etc. that need to be kept confidential. The opinions, data and methods of the person being reviewed are not allowed to be used, shared or discussed with others outside the peer review process. The opinions, data and methods of the person being reviewed are not allowed to be used without permission. Reviewers are not allowed to ask the person being reviewed to cite their own literature.
7. When using generative artificial intelligence in review activities, the consent of the organizer of the review activity should be obtained in advance. The review content should be prevented from being leaked during the

operation. If information leakage occurs, necessary remedial measures should be taken in a timely manner.

8. If during the review process, any violation of scientific research integrity, science and technology ethics, science and technology safety and confidentiality is discovered or has legitimate reasons, it should be reported to the organizer of the review activity in a timely manner. (2) Reviewee

1. Ensure that the materials provided are true, reliable, and accurate, and clearly indicate the source or attribution of all research results, and do not include the research results of other projects or other personnel without explanation.

2. If you believe that a reviewer has a conflict of interest with you, you should report it to the organizer of the review activity in accordance with the procedures, request the reviewer to recuse yourself, and provide sufficient and reliable reasons. 3. Do not interfere with the review process, do not contact privately or bribe or threaten reviewers or review activity organizers.

4. If you have any objection to the evaluation results, you should submit a review application in accordance with relevant procedures. Do not threaten, attack or retaliate against reviewers or review organizers. (3) Review activity organizers 1. Develop scientific, fair and transparent review rules and procedures, and establish reviewer selection, avoidance, work supervision and credit evaluation systems. 2. Strictly perform duties and promptly discover and control possible conflicts of interest. Maintain the independence of review and do not interfere with the legitimate review of reviewers by the will or authority of the organizer or individual.

3. Comply with confidentiality requirements and shall not disclose information such as the list of review objects or reviewers, review opinions, review results, etc. in violation of regulations.

4. Any scientific research dishonesty discovered during the review will be dealt with in accordance with relevant regulations. 7. Ethical review

Science and technology ethics are the values and behavioral norms that should be followed when carrying out scientific research and technological development and other scientific and technological activities. Scientific and technological activities should be carried out in accordance with the principles of "enhancing human welfare, respecting the right to life, adhering to fairness and justice, reasonably controlling risks, and maintaining openness and transparency", and conduct scientific and technological ethical risk assessments or scientific and technological ethics reviews in accordance with regulations. Behaviors that violate the requirements of scientific and

technological ethics should be proactively reported, resolutely resisted, and severely investigated and punished.

(1) Scientific researchers

1. Students should learn science and technology ethics knowledge and relevant management regulations, improve their awareness of science and technology ethics, strictly abide by science and technology ethics norms, and actively participate in science and technology ethics governance.
2. Scientific and technological activities involving human research participants, experimental animals, and scientific and technological activities that do not directly involve humans or animals but may bring ethical risks and challenges in aspects such as life and health, ecological environment, public order, and sustainable development should be carried out. Conduct scientific and technological ethics review in accordance with regulations. Research can only be carried out after obtaining approval, and must not exceed the scope of the approved implementation plan for scientific and technological activities.
3. Research participants should be selected fairly and reasonably to ensure the scientificity, rationality, appropriateness and fairness of their inclusion and exclusion criteria. Research participants shall not be recruited through inducement, coercion, deception, or other improper means. Special protection should be given to research participants involving specific groups such as children, pregnant women, the elderly, people with intellectual disabilities, and people with mental disorders; research involving fertilized eggs, embryos, fetuses, or those that may be affected by assisted reproductive technology should be proactively protected Detailed description.
4. Research participants or their guardians should be clearly informed of all relevant matters and rights they should enjoy, obtain informed consent, ensure that the informed consent process is standardized, and strictly implement agreements or agreements with research participants or their guardians.
5. The collection, storage, use and disposal of biological samples must comply with relevant laws and regulations, and the processing of personal privacy data, biometric information, etc. must comply with relevant regulations on personal information protection.
6. It should be ensured that research is conducted by personnel with relevant qualifications to ensure that research risks are minimized and unnecessary harm to research participants is avoided. If there is reason to believe that death or disability or other injuries will result before the start of the test, the test shall not be carried out; if there are signs during the test that death or disability or other injuries may result, the test must be stopped immediately.

7. To carry out scientific and technological activities involving experimental animals, the use of experimental animals should comply with the principles of substitution, reduction, and optimization, ensuring that the sources of experimental animals are legal and reasonable, and that technical operation requirements such as breeding, use, and disposal comply with animal welfare standards. 8. When carrying out scientific and technological activities involving emergencies such as major public incidents, the scientific and technological ethics emergency review procedures and related requirements shall be observed. Emergency situations shall not be used as an excuse to avoid scientific and technological ethics review or lower the standards of scientific and technological ethics review. 9. To carry out scientific and technological activities that are included in the management of the ethical review review list, after passing the preliminary review by the Science and Technology Ethics (Review) Committee, the scientific research unit should be reported to the local or relevant industry authorities to organize expert review through the scientific research unit. 10. International cooperative scientific and technological activities that require scientific and technological ethics review must pass the scientific and technological ethics review stipulated by the countries where the cooperation parties are located before they can be carried out.

(2) Scientific research units

1. Fulfill the main responsibilities of science and technology ethics management, strengthen science and technology ethics supervision and risk monitoring throughout the entire process of science and technology activities, and actively study and judge and resolve science and technology ethics risks in a timely manner. When carrying out scientific and technological activities that are included in the ethical review review list management, dynamic tracking and ethical risk prevention and control should be strengthened.
2. Establish a science and technology ethics (review) committee in accordance with the actual situation of the unit and provide it with necessary staff, office space, funding and other conditions to perform its duties, and take effective measures to ensure that it independently carries out science and technology ethics review work. 3. Improve the unit's science and technology ethics review supervision mechanism and review quality control, supervision and evaluation mechanism, guide the science and technology ethics (review) committee to formulate charters, and establish system specifications, work procedures and conflict of interest management mechanisms for review, supervision, confidentiality management, file management, etc. , ensuring that science and technology ethics review is compliant, transparent, and traceable. 4. Regularly carry out science

and technology ethics education and training to improve the ethical awareness and risk prevention and control capabilities of scientific researchers.

4 . Academic exchanges and cooperation

Scientific researchers are encouraged to fully exchange academic opinions, research ideas and research findings, and openly share research data and research results in accordance with relevant requirements. When conducting research together, we should strengthen understanding, mutual respect, and promote mutual trust, conscientiously fulfill our responsibilities, and fulfill our commitments and obligations. Through full consultation, the purpose of cooperation, project indicators, expected outputs, rights and responsibilities of all parties, the ownership and use of data and intellectual property rights, and the basis for measuring the contribution of results will be clarified.

1. When conducting academic exchanges, academic democracy should be promoted, originality should be respected, and openness and transparency should be adhered to. You must not use your authority, status, or resources at your disposal to suppress the academic opinions of others.

2. When conducting academic criticism or responding to criticism and doubts from others, rational questioning and criticism should be carried out in a scientific spirit and professional attitude, and the influence of non-academic factors such as personal grievances and conflicts of interest should be excluded. It is not appropriate to make excessive remarks in public and not to easily resort to public opinion or use of online public opinion to coerce academic discussions. Personal attacks and retaliation are not allowed.

3. All parties involved in collaborative research should agree in advance on matters such as rights and obligations, responsibilities and division of labor, allocation of funds, publication and signature of results, ownership of research data and results, intellectual property arrangements and dispute settlement mechanisms through agreements and other forms. The costs and benefits of research should be reasonably allocated among the collaborating parties.

4. All parties involved in multidisciplinary or interdisciplinary collaborative research should understand the research norms and practices of the relevant disciplines. Differences between different disciplines should be negotiated and reached in advance to ensure compliance.

5. In collaborative research, data sources should be compliant and data quality should be ensured, and data from partners should be verified when necessary. without violating relevant laws

5. In collaborative research, data sources should be compliant and data quality should be ensured, and data from partners should be verified when necessary. Under the principle of not violating relevant laws, regulations and confidentiality provisions, relevant research data and research results should be disclosed to collaborators as agreed. 6. In international collaborative research, the relevant regulations on scientific research management and supervision of our country and the country or region where the partner is located should be strictly observed. If you discover or have legitimate reasons to suspect that a partner has engaged in scientific research dishonesty or violated scientific and technological ethics, you should immediately notify the partner and suspend or terminate the cooperation if necessary.

7. If relevant data needs to be provided in foreign scientific and technological exchanges and cooperation, the approval procedures must be completed in accordance with relevant requirements, and relevant laws and regulations on scientific and technological confidentiality and the publication of specific research results must be strictly observed.

9. Intellectual Property Protection

When carrying out scientific and technological activities, one should respect the intellectual property rights of others, abide by my country's intellectual property laws and regulations and relevant international conventions, and strengthen the intellectual property protection, management and application of research results.

(1) Scientific researchers

1. Before publishing research results, the intellectual property protection plan for the research results should be fully considered and reasonable and appropriate intellectual property protection methods should be adopted. If patent protection is adopted, the patent application shall truthfully provide materials in accordance with relevant regulations, and shall not fabricate, forge, or alter the content of the proposed application, experimental data, or exaggerate the technical effects, and shall not plagiarize, simply replace, or piece together existing technology or existing designs.

2. Comply with the laws, regulations and institutional norms of the state and the scientific research unit regarding the ownership of intellectual property rights and the distribution of interests in research results, and promptly and proactively disclose job research results to the scientific research unit where the researcher works.

3. Respect the intellectual property rights of others and do not infringe upon the legitimate rights and interests of intellectual property rights holders. Improve the ability to protect intellectual

property rights, take effective measures to protect the intellectual property rights of important research results, properly handle intellectual property disputes, and safeguard their legitimate rights and interests in accordance with the law.

(2) Scientific research units

1. Improve the intellectual property management system and explore the establishment of specialized internal institutions for the transfer and transformation of intellectual property rights.
 2. Improve the intellectual property management system, clarify the ownership of intellectual property rights and benefit distribution mechanisms for research results, and encourage the creation of intellectual property. Actively promote the transformation and application of research results and safeguard the legitimate rights and interests of scientific researchers.
 3. Strengthen intellectual property risk management, improve scientific researchers' awareness of intellectual property risk prevention, and carry out regular intellectual property risk warnings and infringement risk investigations.
 4. Conduct regular intellectual property training to improve intellectual property management service capabilities and create a good atmosphere that respects creation and intellectual property rights.
10. Training and guidance

Tutors and project leaders should strengthen guidance and supervision of students and research team members. Scientific research units should focus on cultivating scientific researchers' rigorous and serious academic attitudes and truth-seeking and pragmatic scientific spirit, and encourage them to abide by scientific ethics and scientific and technological ethics.

- (1) Graduate tutors and project leaders
1. Practice the requirements of being a teacher and being a role model, give full play to the role of teaching by words and deeds, and regularly provide education and guidance on scientific research integrity, science and technology ethics, etc. to students and research team members.
2. It should ensure that there is sufficient time and energy to guide students and research teams, and provide necessary research resource support.

3. Understand and supervise the daily scientific research activities of students and research team members, follow up and guide experimental progress, review experimental records and data, review research manuscripts, sign signatures of important papers and other research results, research data authenticity, and experimental repeatability etc. for integrity review and academic control. Students and research team members shall not be required or acquiesced in any way to commit scientific research dishonesty or other violations of scientific and technological activities.
4. Adhere to academic democracy and respect the academic opinions and reasonable requirements

of students and research team members for research work. The legitimate rights and interests of students and research team members shall not be infringed upon in aspects such as authorship of results and ownership of intellectual property rights. (2) Students and research team members 1. Invest sufficient time and energy to complete the research tasks assigned by the mentor or project leader, and respect the training and dedication of the scientific research unit, mentor, and project leader. 2. Comply with scientific research management regulations and related requirements, report research progress to the instructor and project leader in a timely manner, collect and save experimental records, data, etc. in accordance with regulations to ensure that the research process is authentic, transparent, and traceable.

3. The announcement, publication or transfer of research results obtained using the research funds, experimental equipment, data materials, etc. of the scientific research unit or team project must comply with relevant regulations. 4. Before leaving school after graduation or leaving the scientific research unit or research team, all original data, pictures, experimental records, samples and other scientific research materials should be returned in accordance with regulations. They are not allowed to take them away without permission. Regarding the permission to use data obtained or accessed from the original unit, the relevant regulations or prior agreement of the original unit shall be followed.

(3) Scientific research units

1. Qualified scientific research units should incorporate scientific research integrity, science and technology ethics, etc. into the curriculum system and allocate corresponding teachers.
2. Establish and improve the education and training system for scientific research integrity and science and technology ethics, carry out scientific research integrity and science and technology ethics training at important points such as enrollment, professional title promotion, and participation in projects, strengthen daily education guidance, and provide scientific research for students and researchers in need. Provide consultation and guidance on integrity and science and technology ethics, and provide timely reminders and conversations to personnel with tendencies and emerging problems.
3. Encourage the development of popular science propaganda on scientific research integrity and science and technology ethics for the public, and guide the public to rationally view the issues of scientific research integrity and science and technology ethics in the development of science and technology.

5. Supervision and management

Adhere to both prevention and punishment, self-discipline and supervision, and prevent and punish irregularities such as dishonesty in scientific research and violations of scientific and technological ethics. Scientific research units should effectively fulfill their main responsibilities, scientific research funding agencies should strengthen supervision of funded projects, the scientific community should play a role in self-discipline and self-purification, and scientific researchers should stick to the bottom line and consciously practice good academic styles.

(1) Scientific research units

1. Establish and improve internal management systems and working mechanisms, and incorporate scientific research integrity, scientific and technological ethics, etc. into normal management. For those who violate the regulations on project application and implementation, fund use, review and evaluation, etc., as well as the requirements for scientific research integrity and scientific and technological ethics, we must dare to expose shortcomings, not accommodate or cover up, seriously investigate and deal with them, and publicly expose them.
2. Organize or entrust a third-party organization to conduct a comprehensive verification of important academic papers and other research results of the unit's scientific researchers. The verification work should be carried out continuously on a 3-5 year cycle.
3. Equip corresponding full-time and part-time personnel to carry out regular work such as monitoring of paper retractions, verification of original experimental data, and review of scientific research integrity involving scientific researchers of the unit.

(2) Scientific research funding agencies

1. Embed scientific research integrity, science and technology ethics and other relevant requirements into the entire management process of project guide preparation, review and project establishment, process management, project completion and acceptance, supervision and evaluation, etc., establish and improve prevention, supervision and investigation and punishment mechanisms, and carry out investigation and punishment according to regulations.
2. Before approval of a funded project, the scientific research integrity status of the project applicant, project leader and relevant personnel should be reviewed. If the project is included in the database of

serious breaches of trust in scientific research integrity and is within the processing period, a "one-vote veto" will be implemented. 3. Designate internal institutions or entrust third-party institutions to supervise and evaluate the implementation of funded projects, and take corresponding measures such as terminating projects, suspending allocations or recovering research funds in accordance with regulations.

(3) Scientific researchers

1. If you discover or have reasonable grounds to suspect that others have violated scientific research integrity standards or violated scientific and technological ethical requirements, you should truthfully report it to the relevant institutions in accordance with relevant regulations. 2. Students should consciously accept academic supervision and proactively cooperate with investigations into scientific research dishonesty and project management inspections. 3. When participating as an expert in supervision and management activities such as investigations into scientific research dishonesty, you should proactively disclose possible conflicts of interest and recuse yourself in accordance with regulations. (4) Science and technology social groups

Scientific and technological social groups such as societies, associations, and research associations must actively play a role in formulating codes of conduct for scientific research activities, providing education and guidance on integrity, and investigating and identifying untrustworthy conduct in scientific research in their respective fields, so as to promote scientific researchers to conduct responsible research and achieve self-regulation., self-management, self-purification. 1. Self-discipline conventions and professional ethics guidelines for scientific research activities in this field should be formulated and improved based on actual conditions, and education on professional ethics and academic style should be regularly carried out to play the role of self-discipline and self-purification. 2. Actively participate in the formulation of relevant institutional norms such as the Code of Conduct for Scientific Research Activities and the Standards for Identification of Untrustworthy Conduct in Scientific Research, and regard compliance with scientific research integrity and scientific and technological ethics as a guide for membership development, award selection, talent recommendation, academician selection, commissioned project evaluation, and youth support projects. and other important conditions for science and technology evaluation activities. Members who violate scientific research integrity standards and scientific and technological ethics requirements will be punished accordingly. 3. Take the initiative

to accept the entrustment of the competent department or relevant units and organize and carry out academic investigation activities related to untrustworthy conduct in scientific research.

4. Set up academic criticism topics, special sections, etc. in hosted conferences and journals, advocate serious academic discussions and comments, encourage young researchers to boldly put forward their own academic opinions, actively communicate and dialogue with academic authorities, and create a free, open, and An equal academic ecosystem.

59. The Australian Framework for Generative Artificial Intelligence (AI) in Schools

The Australian Framework for Generative Artificial Intelligence (AI) in Schools (the Framework) seeks to guide the responsible and ethical use of generative AI tools in ways that benefit students, schools and society. What is Generative AI? Generative AI can generate new content such as text, images, audio, and video that resembles what humans can produce. It is effective at recognising patterns (in video, audio, text or images) and emulating them when tasked with producing something.

Purpose and Audience The purpose of the Framework is to provide guidance on understanding, using and responding to generative AI in Australian school-based education. It supports policy makers, school leaders, teachers, support staff, parents and students. It does not address other forms of artificial intelligence, including predictive AI. The Framework is aspirational in nature, defining what safe, ethical and responsible use of generative AI should look like to support better school outcomes. The Framework's Principles and Guiding Statements are designed to help jurisdictions and sectors align existing approaches while also supporting the development of future work.

Design of the Framework The Framework is designed to help Australian school communities (students, teachers, staff, parents and carers) support:

- **Education outcomes:** The Framework aims to recognise how the appropriate use of generative AI tools can enhance teaching and learning outcomes for all members of Australian school communities.
- **Ethical practices:** The Framework aims to achieve the safe, responsible and ethical use of generative AI tools in Australian schools.
- **Equity and inclusion:** The Framework aims to ensure that generative AI tools are used in ways that are fair, accessible and inclusive of all Australian school communities. These goals are the basis of the Framework's 6 Principles and 25 Guiding Statements. Figure 1 provides a high-level illustration of the Framework, highlighting the interconnectedness of the Principles.

Opportunities and Risks Generative AI technology has great potential to assist teaching and learning and reduce administrative workload in Australian schools. The growing accessibility and sophistication of generative AI tools provides opportunities to develop human-like generated text and rich multimedia content in a way that has not previously been possible. To fully harness the potential of high quality and safe generative AI, schools will need to be supported in understanding and appropriately managing a range of privacy, security and ethical considerations. Risk management should also be appropriate for the potential consequences. These consequences include the potential for errors and algorithmic bias in generative AI content; the misuse of personal or confidential information; and the use of generative AI for inappropriate purposes, such as to discriminate against individuals or groups, or to undermine the integrity of student assessments. Appropriate and proportionate risk management will require robust guidance and policy, which the Framework aims to support.

About the Framework In February 2023, Education Ministers agreed that responding to the risks and harnessing opportunities for Australian schools and students arising from generative AI technologies is a national education priority. Ministers agreed to develop an evidence-informed, best practice framework for Australian schools. The Framework was developed in consultation with unions, teachers, students, industry, academics, and parent and school representative bodies from all sectors. It was developed by the National AI in Schools Taskforce comprised of representatives from all jurisdictions, school sectors, and the national agencies - Educational Services Australia (ESA), Australian Curriculum, Assessment and Reporting Authority (ACARA), Australian Institute for Teaching and School Leadership (AITSL), and Australian Education Research Organisation (AERO). The Framework aligns to existing national policies and goals. It complements Australia's Artificial Intelligence Ethics Framework, and the two key goals of The Alice Springs (Mparntwe) Education Declaration: promoting equity and excellence in education; and enabling all young Australians to become confident and creative individuals, successful lifelong learners, and active and informed members of the community. Additionally, the Framework aligns to the United Nations sustainable development goal #4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. It also aligns with the Safer Technologies for Schools (ST4S) initiative which aims to enhance the security, privacy and online safety of software services and applications commonly used in Australian schools. Any policy development or amendments made to align with the Framework should also

consider the Australian Professional Standards for Teachers (APST), Privacy Act 1988, Copyright Act 1968, the eSafety Commissioner's Best Practice Framework for Online Safety Education, and Australia's human rights protections. Where applicable, policies developed in alignment with the Framework must also meet existing government commitments to incorporate Indigenous community partnerships and expert reviews. Policies must also meet government obligations to ensure Indigenous communities have access to locally relevant data and information as per Closing the Gap Priority Reform 4.

Review The Framework will be reviewed by Education Ministers within 12 months of publication and every 12 months thereafter to accommodate the fast-moving pace of technological development in generative AI. Education Ministers may determine to review the Framework more frequently at their discretion.

Teaching and Learning Generative

AI tools are used to support and enhance teaching and learning. 1.1 Impact: generative AI tools are used in ways that enhance and support teaching, school administration, and student learning. 1.2 Instruction: schools engage students in learning about generative AI tools and how they work, including their potential limitations and biases, and deepen this learning as student usage increases. 1.3 Teacher expertise: generative AI tools are used in ways that support teacher expertise, and teachers are recognised and respected as the subject matter experts within the classroom. 1.4 Critical thinking: generative AI tools are used in ways that support and enhance critical thinking and creativity, rather than restrict human thought and experience. 1.5 Learning design: work designed for students, including assessments, clearly outlines how generative AI tools should or should not be used and allows for a clear and unbiased evaluation of student ability. 1.6 Academic integrity: students are supported to use generative AI tools ethically in their schoolwork, including by ensuring appropriate attribution.

Human and Social Wellbeing Generative

AI tools are used to benefit all members of the school community. 2.1 Wellbeing: generative AI tools are used in ways that do not harm the wellbeing and safety of any member of the school community. 2.2 Diversity of perspectives: generative AI tools are used in ways that expose users to diverse ideas and perspectives and avoid the reinforcement of biases. 2.3 Human rights: generative AI tools are used in ways that respect human and worker rights, including individual autonomy and dignity.

Transparency School communities understand how generative

AI tools work, how they can be used, and when and how these tools are impacting them. 3.1 Information and support: teachers, students, staff, parents and carers have access to clear and appropriate information and guidance about generative AI. 3.2 Disclosure: school communities are appropriately informed when generative AI tools are used in ways that impact them. 3.3 Explainability: vendors ensure that end users broadly understand the methods used by generative AI tools and their potential biases.

Fairness Generative

AI tools are used in ways that are accessible, fair, and respectful. 4.1 Accessibility and inclusivity: generative AI tools are used in ways that enhance opportunities, and are inclusive, accessible, and equitable for people with disability and from diverse backgrounds. 4.2 Equity and access: regional, rural and remote communities are considered when implementing generative AI. 4.3 Non-discrimination: generative AI tools are used in ways that support inclusivity, minimising opportunities for, and countering, unfair discrimination against individuals, communities, or groups. 4.4 Cultural and intellectual property: generative AI tools are used in ways that respect the cultural rights of various cultural groups, including Indigenous Cultural and Intellectual Property (ICIP) rights.

Accountability Generative

AI tools are used in ways that are open to challenge and retain human agency and accountability for decisions. 5.1 Human responsibility: teachers and school leaders retain control of decision making and remain accountable for decisions that are supported by the use of generative AI tools. 5.2 Reliability: generative AI tools are tested before they are used, and reliably operate in accordance with their intended purpose. 5.3 Monitoring: the impact of generative AI tools on school communities is actively and regularly monitored, and emerging risks and opportunities are identified and managed. 5.4 Contestability: members of school communities that are impacted by generative AI tools are actively informed about, and have opportunities to question, the use or outputs of the tools and any decisions informed by the tools.

Privacy, Security and Safety Students and others using generative

AI tools have their privacy and data protected.

- 6.1 Privacy and data protection: generative AI tools are used in ways that respect and uphold privacy and data rights, comply with Australian law, and avoid the unnecessary collection, limit the retention, prevent further distribution, and prohibit the sale of student data.
- 6.2 Privacy disclosure: school communities are proactively informed about how and what data will be collected, used, and shared while using generative AI tools, and consent is sought where needed.
- 6.3 Protection of student inputs: students, teachers and staff take appropriate care when entering information into generative AI tools which may compromise any individual's data privacy.
- 6.4 Cyber-security and resilience: robust cyber-security measures are implemented to protect the integrity and availability of school infrastructure, generative AI tools, and associated data.
- 6.5 Copyright compliance: when using generative AI tools, schools are aware of, and take measures to comply with, applicable copyright rights and obligations.

60. Advice for students on using generative artificial intelligence in coursework

What is Generative Artificial Intelligence or Gen-AI?

Generative Artificial Intelligence or Gen-AI is technology that is used to generate text, images, data, and videos in response to specific prompts. Gen-AI uses machine learning to do this.

A prompt is a descriptive phrase or sentence that directs the Gen-AI to create content.

Examples of Gen-AI

ChatGPT, a text generator

DALL-E, an image generator

[Microsoft Copilot](#), AI platform with text and image generator

As digital technologies become increasingly common, Gen-AI will become widely used in all common software systems that you use in your study, work, and personal life.

This means that, while Gen-AI may be a good resource to support your studies, you must be aware of the risks when using this type of technology.

How and when you can use Gen-AI?

It is your responsibility to use Gen-AI in ways that are allowed by your Course Director and the University.

A Course Director is the person who oversees the running of your course. If your course has more than one lecturer, the Course Director will be likely be the primary lecturer listed on your course outline.

As Gen-AI is not allowed in all assessments, it is important to check the guidelines on permitted use of software in assessment activities. If you are unsure if you can use Gen-AI in a particular assessment, please check out Canvas and/or contact your Course Director.

An assessment can be:

Essays

Projects

Presentations

Reports

Practical, aural, or oral work

Written tests done in exam conditions

Ongoing assessment of competence or performance

Written or performance examination usually completed at the end of semester or year

Guidelines on permitted use of software in assessment activities

If Gen-AI is not approved for your assessment, you cannot use it for that assessment.

If Gen-AI is approved for use in your assessment, your Course Director may specify how you are allowed to use Gen-AI.

As mentioned earlier, please check with your Course Director if you are unsure of whether you can use Gen-AI for a particular assessment.

For example, your Course Director might allow you to use Gen-AI to:

Revise

Refine your writing

Perform certain tasks

If you are allowed to use Gen-AI and you choose to use this technology to support your studies, it is your responsibility to ensure that:

Using Gen-AI is allowed for the particular assessment you are using it for

The final work submitted is your work, reflecting your learning and performance

You correctly acknowledge your use of Gen-AI software*

*To correctly acknowledge your use of Gen-AI software, you can check with your Course Director and/or Canvas for specific instructions on formatting this acknowledgement.

For additional information on how to reference AI-generated content in your writing, please visit [QuickCite](#).

If you use Gen-AI in ways that are not allowed by your Course Director, you may be in breach of the [Student Academic Conduct Statute](#).

Risks to using Gen-AI

There are risks to using Gen-AI to support your studies. It is important to:

Think about and understand what the risks of using Gen-AI are

Critically consider and evaluate what the Gen-AI produces when you use them

Remember that there are differences between different Gen-AI technologies

Accuracy	<p>Even though the information presented sounds convincing, content that is generated by Gen-AI technologies can be:</p> <ul style="list-style-type: none">Non-factualInaccurateOut-of-date <p>You should not use information generated by Gen-AI as your primary and/or only source. You must fact-check all Gen-AI outputs using other reliable sources.</p>
Bias	<p>Content produced by Gen-AI can reflect any biases found within the data sources that the Gen-AI uses that include, but are not limited to:</p> <ul style="list-style-type: none">Discrimination of marginalised groupsUnder-representation of marginalised groups <p>You should critically review any output generated by Gen-AI with the potential of biases in mind.</p>

Quality	<p>Content produced by Gen-AI may lack originality and you are given may not have the tone that you are trying to achieve for your assessment.</p> <p>Additionally, the content may not be written in a way that is appropriate for your assessment.</p> <p>For these reasons, remember that the content generated by Gen-AI may not be of a high standard.</p>
Privacy	<p>As Gen-AI learns from everything that you put into it, please make sure to not enter any:</p> <ul style="list-style-type: none"> Personal details Confidential information Confidential data <p>This is because the Gen-AI you use will remember what this information is and may use it in other generated responses.</p>

How Gen-AI may support your studies

You may be allowed to use Gen-AI to support you in your studies. Your Course Director will specify how you can use Gen-AI in their course.

These are some of the ways that Gen-AI may support your studies:

Help you to revise your course content, such as through generating practice questions or flashcards

Get a general overview of your topic before you begin researching in more detail

Help you to improve your writing skills, including grammar, use of vocabulary, and sentence structure

Explain solutions to problems in another way to help you understand how a particular answer was arrived at, such as in mathematics

Help you to analyse data and/or organise information

Misusing Gen-AI

There are many ways that Gen-AI can be misused in your studies, including:

- Using Gen-AI in assessments that you have been told the use of Gen-AI is not allowed
- Using Gen-AI in ways not allowed by your Course Director
- Using Gen-AI to write or develop your assessments for you. All work that you submit must be your own work and representative of your capabilities. This includes where your Course Director allows you to use Gen-AI to support your learning
- Using Gen-AI where it is allowed and not correctly acknowledging its use

If you use Gen-AI in any of the mentioned ways, you may be in breach of the [Student Academic Conduct Statute](#).

Microsoft Copilot

The University of Auckland is pleased to announce that a secured version of Microsoft Copilot is now available for students. This tool is a powerful Gen-AI platform currently based on GPT-4 and DALL-E 3 that can assist and enhance your learning experience. This version of Copilot provides a secure environment for your work, the information you enter into Copilot is protected.

Accessing Microsoft Copilot

To access Microsoft Copilot, go to <https://copilot.microsoft.com> with your internet browser. Ensure you log in using your University of Auckland account. Upon successful login, a green “Protected” icon should appear on the top right of the Copilot interface, indicating that you are in the secured version.

Responsible Usage

While Microsoft Copilot is a powerful tool, it's important to use it responsibly. The provision of this tool by the University of Auckland does not constitute an endorsement to use Gen-AI in all situations. Always refer to the guidelines on responsible Gen-AI usage provided in other sections of this webpage. Remember, the goal is to enhance your learning, not to replace it.

We encourage you to explore this tool and make the most of its capabilities responsibly while adhering to the principles of academic integrity.

Additional information

To learn more about academic integrity at the University of Auckland and effective study skills to support your learning, please check out:

If you're not sure about using Gen-AI

Course-specific questions

If you have questions about using Gen-AI in your courses, get in touch with your Course Director and/or tutor. They can help answer questions about Gen-AI use in your specific courses.

You can find your Course Director's details on [Canvas](#).

Questions about University policy and procedures

If you have general questions about University policy and/or procedures around using Gen-AI, you can email the Academic Quality Office.

[**Academic Quality Office**](#)

Academic skills support

If you are looking for support in developing your academic skills, you can get in touch with the Learning Advisers at Libraries and Learning Services.

[**Learning advisers**](#)

61. Use of generative AI tools to support learning

Overview

Part of what a university education teaches is certain academic skills, such as assimilating information, constructing an evidence-based argument and expressing your thoughts in clear, coherent prose.

AI tools cannot replace human critical thinking or the development of scholarly evidence-based arguments and subject knowledge that forms the basis of your university education.

You can make use of generative AI tools (e.g. ChatGPT, Claude, Bing Chat and Google Bard) in developing your academic skills to support your studies. Your ongoing critical appraisal of outputs by reviewing them for accuracy will maximise the potential for AI outputs to be a useful additional tool to support you in your studies.

In some instances academic staff, departments and colleges may give more detailed guidance on how they expect AI tools to be used (or not used) for different tasks or on specific assignments. You should always follow the guidance of your tutors, supervisors and department or faculty.

Ethical use of generative AI tools

Your ethical use of AI tools to support your learning is paramount to ensure you uphold the standards of academic rigour and academic integrity expected of you as a student studying at Oxford.

Students using AI during their studies must learn and practise the same academic skills of note-taking and clear attribution which are safeguards against plagiarism, ensuring clear differentiation of their own work from any text or material derived from any AI tools. Unauthorised use of AI falls under the plagiarism regulations and would be subject to academic penalties in summative assessments.

Where the use of generative AI in preparing work for examination has been authorised by the department, faculty or programme, you should give clear acknowledgment of how it has been used in your work.

The advice in this guide is in line with relevant university policies, which have been updated to provide examples of unauthorised uses of generative AI tools. [Guidance on plagiarism](#) is available from the Study skills and training webpages. The policy on the [Use of Third Party Proof-readers](#) is also relevant.

Five things to think about when using generative AI tools

1. How can generative AI tools be useful in supporting your learning?

2. How can generative AI tools be useful in developing your academic skills?
3. How (and when) can you draw on AI outputs (knowing that different generative AI tools provide different outputs and that the same tool with the same prompts can produce different outputs)?
4. How can you manage the risk of false information and fabrication?
5. How can you ensure you maintain good academic practice?

Six tips to keep in mind when using generative AI tools

1. Always cross-check AI generated outputs against established sources to verify accuracy and identify erroneous information.
2. Give significant contextual information when asking questions or prompts and ask several follow-up questions to refine responses.
3. Use personae in your prompts e.g. “I am an undergraduate student who is revising for a first-year calculus exam”.
4. Give examples of the kind of responses you want.
5. AI tools are not good at calculations so use other established tools, calculators, Excel or Mathematica.
6. Do not share sensitive personal data such as financial details or passwords with AI tools. Avoid sharing your own or others intellectual property such as patents, trademarks, designs, sensitive information, or content created by others into any AI tools.

Five ideas for academic reading

Generative AI tools may be useful in supporting you to develop your academic reading skills. However, generative AI may, in some cases, undermine development of your academic reading skills (e.g., asking an AI tool to summarise an article rather than undertaking the task yourself). Here are five ways you can use AI to support your academic reading, suggestions for how you might provide useful prompts, what some of the limitations are of the AI generated outputs, and how to engage critically with them to augment your learning.

1. When reading a paper ask for a table of key terms or outline key points in the paper. Do this yourself before asking AI to do this and compare your terms or points. The AI tool can help you build a cognitive scaffolding of your reading of a paper but you cannot rely on it, so ensure you read the paper yourself.
2. Ask AI to generate thought-provoking questions based on article content. You can develop your own understanding of an article by answering the questions asked. You could also use the questions to develop your own questions in relation to the article to deepen your learning.
3. Ask AI to translate sections into another language. You can ask AI to translate text into your own language. Be sure to check the accuracy of the translation and that no significant meaning has been lost in translation. This could support your understanding of the article as you review the translation.
4. Compare your own summary of a paper with one written by AI. AI can be a useful tool for providing a summary and supporting your reading of academic papers. Comparing your own understanding of the paper with an AI output can be a useful approach to developing your critical reading skills – both by recognising things you may have missed, and by giving you an opportunity to critique the AI output.
5. Critically review all AI outputs. Ensure you critically review all AI outputs for accuracy to support your learning, and verify the

outputs against other sources, e.g. can you locate all references and are they accurate to the text, and are the definitions correct?

Suggestions of prompts to try:

You are an expert in [subject] and an educator who is good at giving great explanations to beginners. Make a table of the key concepts needed to understand the content of this paper.

Give me a list of 20 key terms in this paper and break it into five categories.

Rephrase this definition as a list of bullet points to help me understand it step by step.

Make a list of propositions in this text in the format “X is a type of Y”, “W is caused by X”, “A explains B”. Put it into a table with three columns.

Rephrase this sentence in simpler language that a non-expert can understand. You can break it into multiple sentences, if needed.

Five ideas for academic writing and presentation skills

Generative AI tools can be useful in developing your academic writing skills and providing initial feedback on them, translating between different styles and critiquing writing. AI tools cannot replace the need for you to develop these skills through teaching and independent learning.

Here are five ways you can use AI to develop and get feedback on your academic writing, suggestions for how you might provide useful prompts, what some of the limitations are of the AI generated outputs, and how to engage critically with them to augment your learning.

1. Examples of writing in different styles and genres. AI can provide examples of many different written outputs. You may find this useful in identifying different styles of writing, appropriate to different tasks.

2. Feedback on your writing. AI can provide rapid feedback on your writing, and this can be helpful for improving it, e.g. in relation to grammar and structure. AI cannot provide feedback nuanced to the rigour and expectations of academic writing in your specific subject so it cannot replace tutors' feedback grounded in the context of your discipline. Be sure that you proof-read your own work, as this is an essential skill in academic writing, and do not use generative AI tools to make material changes to work in draft.
3. Help you get started in writing. AI can be useful in overcoming writer's block by providing some inspiration or points to consider when you are about to start writing.
4. Suggesting ideas for graphics, images and visuals. AI can be useful in suggesting how you might present information in graphics, images and visuals to move beyond text-based presentation of information.
5. Critically review all AI outputs. Ensure you critically review all AI outputs for accuracy and verify the outputs against other sources, e.g. can you locate all references and are they accurate to the text, and are the definitions accurate?

Suggestions of prompts to try:

Here is a paragraph I wrote. Rewrite it as a list of bullet points with the key arguments I made.

I am a student of [subject]. I wrote this as an example of academic writing. Give me feedback on where I could improve. Focus on clarity, academic language and grammar.

Suggest some practice exercises I can use to improve my weak points.

Five ideas for supporting your learning

Generative AI tools can be useful in supporting your academic studies.

Here are five ways you can use AI to support your studies, suggestions for how you might provide useful prompts and what some of the limitations are of the AI generated outputs. Be sure to verify any AI outputs against other established sources to ensure their accuracy, and cross-check to confirm your understanding.

1. Preparing for lectures. You can ask for key concepts related to the topic of the lecture and use the lecture to compare this with your initial understanding.
2. Engaging with new or complex topics. You can ask AI for explanations to help you develop a better understanding for yourself. For example, you can ask for alternative explanations of a topic or analogies from different perspectives.
3. Organising your notes. Generative AI tools can be used to convert your notes into structures that are easier for you to review. You can ask for a table of key concepts, facts or figures organised by different categories. Timelines, workflow summaries, outlines you can convert into mind maps or mnemonic devices can be generated to help you remember facts. Remember, even if AI looks like it is copying things from your text, it is actually generating the new version from scratch using your text as context, so you must review all outputs for accuracy.
4. Enhancing your language study. You can use AI to improve your language learning. You can ask to have a conversation at a certain level on a specific subject, receive feedback on your conversation, or generate sample texts for practice. Always ensure you cross-check the outputs against other sources as outputs in some languages may contain basic grammatical errors. If working with

long texts, for languages other than English, the volume of text you can translate or work with will be much shorter. Some languages with non-Latin alphabet (like Arabic) may only be able to fit 1/5th or even less than a language like French. Many languages not well represented on the internet may only work very partially with generative AI.

5. Developing your coding skills. You can use AI to develop your coding skills. Write code as you do normally – writing in (functional) chunks and testing the chunks for the right functionality yourself. You could then compare these chunks with a given AI tool. You may find it useful to ask your tutor which AI tools are more appropriate for coding. As with all AI outputs it is not perfect, and any code needs to be run to check that it functions as expected. Common errors are: the code may produce an error message, the description of the code or suggestions for improvement are inaccurate, the code is completely incorrect, or uses features not present in the language or the code generates references to outdated or non-existent libraries. It would not be appropriate to use AI to write a code, or to support the writing of a code, on which your coding skills will be assessed.

Suggestions of prompts to try:

I'm going to a lecture on [subject]. Give me a list of 20 key concepts I need to be prepared for. Break them into five categories and explain how they relate to each other. You are an expert in [subject] and always explain things in a way that is easy to understand for a student at [level]. Make a table of these terms with four columns. 1. Term, 2. Definition, 3. Category, 4. Related terms. Make a series of statements using these terms that will reveal the

relationships between them. For example, working memory is a type of

As an expert in [subject]. Here's an explanation I was given of [concept]. Can you suggest some alternative explanations of [concept] to help me learn. I am a [level] student at university and have already taken [classes in...].

You are an expert in [subject]. Here is a definition I was given of [term]. Can you give me five concrete examples that illustrate the definition. I want to be able to apply them in practice. For each example, specify exactly how it illustrates the definition.

Here are my notes from class. Make me flashcards to help me study. Each flashcard should have the term on one side, and a brief explanation on the other. Use my explanations where available. If you have suggestions for alternatives, put them in italics so that I know what is in my words.

You are a shop assistant in a bakery in Germany. I am a beginner student and I've come in to buy some bread. Let's have a conversation in German. You start by greeting me and asking me what I want. You will pause and give me a chance to answer. Use simple language and always respond even if I say something using the wrong term. Do not use any English until I tell you to. Then make a table of all the things I said with three columns: 1. What I said, 2. What I should have said in correct German. 3. What error I made and how to correct it.

I want to write code to do [describe task]. As an experienced programmer and a coach of new coders, suggest the best way to help me get started. Suggest the best language, coding environment, and dependencies. My level is [describe level].

Five tips for selecting the right generative AI tools for the task

There is a huge range of generative AI tools available, e.g. Microsoft Bing Chat, ChatGPT, Google Bard, Claude. Some of the tools are freely available and some with costs that bring extra features such as being able to work with large volumes of text.

1. You will find it useful to try a few different AI tools and be aware that different tools will give different outputs using the same prompts and a tool will also generate different outputs to the same prompt.
2. You will get different responses using the same prompts from the same AI tool. AI outputs are not repeatable, and all tools can generate outputs that can contain inaccuracies and fabrications.
3. You could spend a lot of your study time trying out different AI tools. Be careful to manage your exploration of tools alongside managing your time.
4. AI tools may draw on data that can be months or years out of date and whilst outputs seem plausible, they may contain errors and/or reflect biases from the original data, e.g. Western perspectives are overly represented.
5. AI tools will not replace the need for you to develop your own knowledge and skills as an independent learner.

Further resources

[Guidance on plagiarism](#)

[Policy on the Use of Third Party Proof-readers](#)

62. Artificial intelligence and teaching, learning, and assessment

Artificial intelligence and teaching, learning and assessment

Cambridge recently collected feedback from our schools on the use of artificial intelligence in education. View the [results on the blog](#).

At the end of 2022, [OpenAI](#) released a conversation based artificial intelligence (AI) tool (or online response generator) – [ChatGPT](#). It was described as the fastest growing technology platform ever, with one million users within five days. The tool has seen significant use within the education community, and has drawn both interest and concern, due to its ability to produce human-like responses to instruction, including homework tasks, assignments and coursework.

Since its release, some schools are concerned students may be using online response generators like ChatGPT, to help with written work, or even submitting work produced by these tools in place of their own. Due to online response generator's ability to produce plausible writing at the level and in the style of school-age learners, we understand teachers are concerned about how to tell when they have been used.

Use of online response generators cannot always be identified by detection tools, which can be unreliable - generating both false negatives and false positives. This can create doubt in teachers' minds and introduce a sense of mistrust into their relationships with students.

To support teachers, we encourage schools to consider the following guidance:

- Review the school's 'academic honesty' and 'approaches to learning' policies, to ensure teachers talk with students about how and when large language models (LLMs) such as ChatGPT and generative AI can be used. We are [not recommending banning the use of these tools](#), but students need to be clear about what they can be used for and when they must not be used. They must also understand how they should be referenced.
- Discuss the strengths and weaknesses of LLMs and the technology that underpins AI tools, such as OpenAI's ChatGPT or Google's Bard. On the one hand, they might be a good way for students to create an essay framework or to help them compile a written response from multiple AI outputs. On the other, these tools can suffer from 'hallucination' and make factual errors that, on the face of it, look highly plausible, so it is a good idea to work with students on how to check for inaccuracies and inherent bias.

- Carefully consider the command words you might want to use with students, as they engage with such tools. Examples could include: prompt, compile, modify, correct, improve upon, generate, challenge - all of which foster higher-order thinking and deeper levels of understanding.
- Encourage teachers to collaborate on the changing nature of assignments and tasks they set for students and to consider ways they could ask for evidence of understanding that is not just based on prose or short text – for example, label a diagram, deliver a presentation, employ the Socratic method, or create a flow chart. Using images alongside thinking routines, such as See, Think, Wonder, can help to both uncover student thinking, as well as to develop it. Also, requiring students to draw on specific perspectives or examples from class in their written work will ensure contextualisation that is harder for generative AI to reproduce.
- Ask students to reference their sources and create bibliographies, as LLMs currently struggle to provide accurate citations – although this seems likely to change soon.

We understand that AI tools to support teachers are also emerging. These tools can help to create handouts, lesson plans and schemes of work, based on instructions provided by the teacher. Tools like this would not only save time and increase productivity, but also support teachers in the process of scoping and sequencing a learning progression – a critical element of scaffolding learning for students.

Ultimately, we recognise that AI is creating a whole new paradigm in teaching, learning and assessment. Cambridge is developing a strong understanding of how this will evolve, and we are committed to playing a key role in supporting schools as they look to embed these new models into education policy and practice.

We have published separate guidance on identifying the use of the online response generators in your students' coursework - [Preventing plagiarism using online response generators](#).

63. Generative AI Guidance

What are Generative AI tools?

Generative AI tools are a type of software that automatically generate content based on questions or prompts input by the user. They include ChatGPT, Bard and Synthesia, among many others. Tools can generate text, code, images, and other types of content. Generative AI tools that deal with text are also called Large Language Models, or LLMs and continuously build a ‘knowledge base’ of information by collecting users’ previous questions or prompts as its base of data. Generative AI tools can be appealing to use in that they can provide or create content quickly but have limitations for use in any academic work in that the outputs may not be wholly reliable.

[An introduction to Generative AI Technology from JISC](#)

Acknowledging and referencing and generative AI tools

Acknowledging

You should include a statement to acknowledge your use of generative AI tools for all assessed work, in accordance with guidelines from your department or course team.

This statement should be written in complete sentences and include the following information:

- Name and version of the generative AI tool used; e.g. ChatGPT-3.5
- Publisher (name of company that provides the AI system); e.g. OpenAI
- URL of the AI tool
- Brief description (single sentence) of the way in which the tool was used
- Confirmation the work is your own

For example:

I acknowledge the use of ChatGPT 3.5 (OpenAI, <https://chat.openai.com/>) to generate an outline for background study. I confirm that no content generated by AI has been presented as my own work. Further requirements may be stipulated for a particular piece of assessed work and must be made clear to students when it is set. Additional requirements may include expanded description in the ‘Acknowledgements’ or ‘Methods’ section, such as:

- If relevant, the prompt(s) used to generate a response in the AI system.
- The date the output was generated.
- The output obtained (e.g. a ‘link to chat’ if ChatGPT, or a compilation of all output generated as an appendix).

- How the output was changed for use or incorporation into a piece of work (e.g. a tracked-changes document or a descriptive paragraph).

[References, citations and avoiding plagiarism](#) by [UCL Library Services](#), used under [CC BY-NC-SA 4.0](#) / edited from original.

Referencing

Referencing guidance is provided for Harvard and Vancouver referencing styles on the following pages:

[Your reference list and bibliography – Harvard](#)

[Your reference list and bibliography – Vancouver](#)

Considerations for use of generative AI tools

There is no agreed AI Literacy framework yet in the Higher Education sector to outline necessary skills and competencies for students and academics when using these tools. Library Services has compiled a list of considerations below for the use of generative AI tools.

These considerations are organised according to the continuum of information use and management the library supports. They are not intended to be exhaustive. For further information consult your course team or tutor and your [Subject Librarian](#). The College [Academic Misconduct Policy and Procedure](#) addresses the use of generative AI tools and academic integrity should be understood to encompass appropriate use of these tools.

Accordion widget

[Collapse all](#)

Finding information

It is acceptable to use a generative AI tool as you would a search engine, for example, to look up background information. Any information later used must be referenced and cited properly. Please refer to the Library's [reference management pages](#).

- Generative AI tools can provide a starting point for research. We do not recommend using generative AI for in-depth research or instead of using credible sources such as academic databases. To identify and use the best academic sources for your work, your [Subject Librarian](#) can provide guidance and advice.

Critically evaluating information

- Generative AI tools' outputs are, fundamentally, not reliable nor trustworthy as pure statements of fact. AI 'hallucinations' occur when factually incorrect information is included in the generative AI tool's outputs. Therefore you cannot trust the tool as the sole source of any facts and it is essential to verify any statement generated by these tools.
- Your subject librarian and the library provide in-depth content and advice and guidance on critical evaluation of sources and the best sources to use for academic work.
- Using generative AI tools to analyse numeric data is not recommended at this stage.

Ethically using information

- It is important to familiarise yourself with your department's current guidance on the use of generative AI tools in support of academic work. [As stated in College guidance](#), your department may choose to invite a random selection of students for 'authenticity interviews' on submitted assessments.
- You may be asked as part of your assignment or assessment to use generative AI tools. We recommend that clarity is sought from your course team or tutor if you are unsure how to do this. Where there is no explicit instruction to use generative AI tools, it would not be considered acceptable to use them to write your assessed work.
- It is important to disclose the use of generative AI tools for academic work according to your department's requirements. You can also refer to the library's guidance on proper citing and referencing on the [reference management pages](#).
- Unless explicitly authorised to use as part of an assessment, the use of generative AI tools to create assessed work can be considered a form of contract cheating, which is addressed on the College [Plagiarism, Academic Integrity & Exam Offences web page](#) as well as within the library's [Plagiarism Awareness](#) courses.
- It is not advisable to add sensitive data (such as student names or other personal data) into generative AI tools, such as ChatGPT, as queries are stored and become part of the training data it draws upon.

Creating and communicating information

- When considering the use of generative AI tools for creating original academic work, ask yourself if it would be acceptable to ask another person to do this work for you. If the answer is no, then it is not acceptable to use these tools.
- Consider that outputs from generative AI tool outputs can be quite generic. You risk losing your own original voice when using it to create written work. Using these tools uncritically would likely result in a text that adopts an academic style superficially but contains little in the way of substantial communication.
- It may be unacceptable to use an essay outline produced by generative AI tools.

64. School Statement on Generative Artificial Intelligence and Education

School statement on Generative Artificial Intelligence and education (updated 10 November 2023)

Academic and professional staff across LSE are committed to providing students with an excellent education and student experience during their time at our School. We are committed to upholding high academic standards and rigour, ensuring the value of our awards lies at the heart of our educational offer.

Alongside this, LSE embraces the exciting opportunities Generative Artificial Intelligence tools (Generative AI) present to enhance teaching, learning and assessment for our diverse community of staff and students. The evolving sophistication and capabilities of Generative AI tools will have a significant impact on various aspects of our students' educational experience, and future career trajectories.

As a result, LSE takes a proactive approach to reassessing teaching and assessment in relation to Generative AI, while emphasising the importance of ethical scholarly practice and upholding the integrity of our assessments. Working in partnership with students is an integral part of LSE's approach, as is our commitment to provide clear and accessible guidance to students on the authorised use of generative AI tools on their programmes and courses.

Over the last year, Vice President and Pro-Vice Chancellor for Education, Professor Emma McCoy, and the Director of the Eden Centre for Education Enhancement, Dr Claire Gordon, have convened a cross-School Working Group on Artificial Intelligence and Education to consider the impact of Generative AI and make recommendations to our community. This group, comprising academic, professional staff and students, has been actively exploring the challenges and the

potential of Generative AI for higher education and our community of students and educators. Several exciting projects, including on student voice and the application of generative AI tools in teaching, learning and assessment, have been established. As technology continues to evolve, the guidance will be adapted accordingly, and regular updates on progress will continue to be shared with our School community

65. Engaging with AI in your education and assessment

Using AI tools in assessment

A three-tiered categorisation of AI use in an assessment for staff to guide students on expectations

Introduction

Staff have questions about how generative AI can be integrated into assessment design in ways that prepare preserving academic integrity. Students report being confused about what constitutes academic integrity in different process, especially when AI tools are becoming so widely available. This guidance will help staff and students to understand. Like many sectors in society, higher education is both challenged by recent developments in AI, and intrigued by them. The use of AI in assessment has generated numerous questions in our sector.

There are no simple answers and our responses will require constant revision as generative AI continues to evolve.

Using AI to support learning

As a general principle, we need to recognise that AI will be used by students at many different stages in their learning.

Our goal, therefore, is to ensure that students are using AI in ways that support their learning, enhance their ability to succeed in their studies, and prepare them to succeed in their future careers. Inappropriate use of AI will undermine all of these benefits and potentially lead to academic misconduct.

A key element of such an approach is communication with students so that they are fully aware of the parameters of AI.

Therefore, the use of generative AI does not automatically constitute academic misconduct. Whether its use is acceptable or not depends on the context and the specific circumstances. This guidance seeks to support staff in considering and clarifying what is, and is not, acceptable. In this section, we will focus on the use of AI in assessment, referring to generative AI technologies.

Note that UCL is not investing in AI detection software. However, staff must not submit student work into such systems without clear consent from the student. It is important to respect the student's intellectual property and personal data rights. It must also be made clear to students that they must not upload any personal data to such systems without appropriate consent. Personal data protection requirements. Approvals and consents are likely to be required. As such personal data usage should be made to [UCL's separate data protection guidance](#).

If AI is misused in assessment, this would be considered under the category of plagiarism or falsification, not conduct.

Keeping up with developments

Given the speed at which generative AI tools are being developed, and with which educational uses of them is emerging, it is important to stay informed about new developments. There may, for example, be some emergent tools that are particularly relevant for your discipline. It is also possible that some AI tools will emerge with some AI tools which mean they are no longer considered to be acceptable to be used in certain ways. UCL's Academic Integrity Policies and Processes are currently being reviewed and updated, and will be available on the Academic Integrity web pages. Information on academic integrity and academic misconduct is available [in the Academic Integrity web pages](#).

Guidance

Colleagues from across UCL have pooled their expertise to offer a broad three-tiered categorisation of AI use in assessments. These categories are designed to help staff to support students in using AI tools to support their learning and development, rather than to rigidly limit their use. The categories are:

- Category 1: AI tools cannot be used*
- Category 2: AI tools can be used, but staff must set clear rules for how they are used
- Category 3: AI tools can be used, but staff must support students to use them effectively

These categories are designed to help staff to support students in using AI tools to support their learning and development, rather than to rigidly limit their use. The categories are:

- Category 1: AI tools cannot be used*
- Category 2: AI tools can be used, but staff must set clear rules for how they are used
- Category 3: AI tools can be used, but staff must support students to use them effectively

Staff should discuss with students the category that their assessments fall into at beginning of each module, and throughout the module. Whilst colleagues may be tempted to ban use of AI tools (Category 1), they might want to consider whether this is appropriate given the nature of the assessment, the purpose of the assessment, and the potential benefits of using AI technology (as opposed to cheating) technology and whether any ban can be enforceable. If you have questions, please discuss with your HEDS Faculty Partner or the HEDS Team.

Where students are using generative AI in assessed work, they should acknowledge how they have used it as part of their work. Please see [students on how to acknowledge the use of generative AI](#).

Students should always be strongly encouraged to take a critical approach to the use of any output from a generative AI tool, and to be aware of the potential for inaccurate and unhelpful outputs.

The categories of assessment

1. AI tools cannot be used*

The purpose and format of these assessments makes it inappropriate or impractical for AI tools to be used. Assessments where the use of AI is wholly inappropriate for the delivery of the specific learning activities or outcomes include those that require students to demonstrate their ability to demonstrate foundation level skills such as remembering, understanding, independently developing critical thinking skills, and applying knowledge to solve problems. These fundamental skills that will be required throughout the programme.

Such assessments are likely to be designed to support the development of knowledge and skills that students will effectively, including with the use of AI tools in other contexts and in future assessments. Discussion with student category (for example, pedagogy, employability, etc).

Examples of assessments where AI might not normally be used could include:

Students believed to have ignored the categorisation will undergo the standard academic misconduct procedure. Note that in UCL's Language and Writing review in [the Academic Manual \(9.2.2b\)](#), it is permissible for a third party to "make changes to content, structure, fluency, presentation, grammar, spelling, punctuation, and language translation." However, "this may reflect changes to content have been made by the reviewer or software or at their recommendation."

***Students with a Summary of Reasonable Adjustments (SORA)**

2. AI tools can be used in an assistive role*

Students are permitted to use AI tools for specific defined processes within the assessment.

AI tools can be utilised to enhance and support the development of specific skills in specific ways, as specified. For instance, students might use AI for tasks such as data analysis, pattern recognition, or generating insights.

Here the tutor should support and guide the students in the use of AI to ensure equity of experience, but the use of AI is inappropriate in some aspects of the assessment where the use of AI is inappropriate.

Examples of where AI might be used in an assistive category include:

3. AI has an integral role

AI can be used as a primary tool throughout the assessment process.

Students will demonstrate their ability to use AI tools effectively and critically to tackle complex problems, make informed decisions, and communicate their findings. The assessment will provide an opportunity to demonstrate effective and responsible use of AI. The tutor should ensure equity of experience.

Examples of where AI tools could be used as an integral part of the assessment include:

UCL guidance on acknowledging use of AI and referencing AI

Generative AI is evolving rapidly and there is not yet consensus on how to acknowledge and reference it. This updated.

Guidance can be found on the [Library Skills](#) pages.

66. AI Guidance for Staff and Students

Generative Artificial Intelligence guidance for students

Guidance and advice for students on the use of Generative Artificial Intelligence (such as ChatGPT) within the University.

The technology, ethics, and use of AI is a fast moving area. This guidance is current as of March 2023 and will be updated as necessary.

University position

There is currently a lot of interest in generative AI systems. ChatGPT (by OpenAI) is just one example, but there are others (such as DALLE-2, CoPilot, and Google Bard). It is an exciting area and naturally we want to explore what it can do and learn how to make use of it.

The University position is not to impose a blanket restriction on the use of generative AI, but rather to:

- Emphasise the expectation that assignments should contain students' own original work;
- Highlight the limitations of generative AI and the dangers of relying on it as a source of information;
- Emphasise the need to acknowledge the use of generative AI where it is (permitted to be) used.

Some assignments may explicitly ask you to work with AI tools and to analyse and critique the content it generates, others may specify that AI tools should not be used, or only used in specific ways. This will depend on the learning objectives for your courses. Please refer to the specific criteria for your assignments and ask your lecturers if in doubt.

Expectation of own original work

All work submitted for assessment should be your own original work. In some cases, you may be asked to sign a declaration of own work. It is not appropriate to misrepresent AI generated content as your own work.

Important note

Be aware that if you use AI tools (such as ChatGPT or others) to generate an assignment (or part of an assignment) and submit this as if it were your own work, this will be regarded as academic misconduct and treated as such.

“Academic misconduct is defined by the University as the use of unfair means in any University assessment. Examples of misconduct include (but are not limited to) plagiarism, self-plagiarism (that is, submitting the same work for credit twice at the same or different institutions), collusion, falsification, cheating (including contract cheating, where a student pays for work to be written or edited by somebody else), deceit, and personation (that is, impersonating another student or allowing another person to impersonate a student in an assessment).” (University of Edinburgh, Academic Misconduct Procedures)

Current limitations of generative AI

Generative AI offers a number of benefits, but it also has its limitations, which you need to aware of.

It is important that you

- Understand the limitations of any AI system you are using;
- Check the factual accuracy of the content it generates;
- Do not rely on AI generated content as a key source - use it in conjunction with other sources.
- Generative AI tools are language machines rather than databases of knowledge – they work by predicting the next plausible word or section of programming code from patterns that have been ‘learnt’ from large data sets;
- AI tools have no understanding of what they generate. A knowledgeable human must check the work (often in iterations);
- The data sets that such tools are learning from are flawed and contain inaccuracies, biases and limitations;

- They generate text that is not always factually correct;
- They can create software/code that has security flaws, bugs, and use illegal libraries or calls – or infringe copyrights;
- Often the code or calculation produced by AI will look plausible but contain errors in detailed working on closer inspection. A human trained in that programming language should fully check any code or calculation produced in this way;
- The data their models are trained on is not up-to-date – they currently have limited or constrained data on the world and events after a certain point (2021 in the case of ChatGPT);
- They can generate offensive content;
- They produce fake citations and references;
- Such systems are amoral - they don't know that it is wrong to generate offensive, inaccurate or misleading content;
- They include hidden plagiarism – meaning that they make use of words and ideas from human authors without referencing them, which we would consider as plagiarism;
- There are risks of copyright infringements on pictures and other copyrighted material.

Important note

Over-reliance on AI tools simply to generate written content, software code or analysis reduces your opportunity to practice and develop key skills (e.g. writing, critical thinking, evaluation, analysis or coding skills). These are all important skills that are valued and required to succeed in and beyond your time at University.

Citing and acknowledging the use of AI

Where the use of AI is permitted in assessed work, it is important to be transparent about the use of such tools and content generated from them.

Content generated from AI is non-recoverable - it can not be retrieved or linked to in the same way that other digital sources can. For this reason, current convention is to cite AI generated content as “personal communication” (because it is based on asking a question or giving a prompt and receiving an answer). This is usually an in-text only citation.

Each reference style (e.g. Harvard, APA) will set out how to do this, so you should consult the guidance for the reference style you are using.

Additionally, if you use any generative AI tool (such as ChatGPT) to help you (e.g. generate ideas or develop a plan), you should still acknowledge how you have used the tool, even if you do not include any AI generated content in your work. You should acknowledge the AI tool used, describe how you used it, and indicate the date you accessed it.

Further guidance

[Further guidance on academic misconduct \(including plagiarism\) and how to avoid it](#)

[Further guidance on general referencing](#)

[Academic Misconduct Investigation Procedures \(859.65 KB PDF\)](#)

67. Generative AI in education

There are many exciting developments happening in the field of Artificial Intelligence (AI). What implications does this have for education at Erasmus University Rotterdam (EUR)? On this page, you will find an overview of relevant information and developments within our university. This information has been gathered from contributions by experts, users, and task forces. AI has been on everyone's mind for a while now, with developments following each other at lightning speed. For this reason, this page will be updated regularly.

Generative AI

Recently, generative AI has received a lot of attention. This is a branch of AI that focuses on developing algorithms and systems capable of creating new content, such as images, music or text. Unlike traditional AI systems that are programmed to perform specific tasks, generative AI algorithms can learn from user input and independently generate content similar to human creations. This is done, for example, through machine learning, deep learning or neural networks. Generative AI has many applications, including the development of chatbots, video games, virtual assistants and more. ChatGPT is now a well-known example of this, but numerous other examples can be found.

Practical information for lecturers

Disclaimer

The information under this heading was collected by the ESE Learning and Innovation Team and the EUR task force AI. Given the rapid pace of developments, we cannot guarantee that information is always fully up-to date. Updates will be made when necessary. If you have any suggestions or updates to this page, please email the Community for Learning & Innovation. For any school-specific questions, please reach out to your [Learning & Innovation team](#)

[Opens external](#)

Generative AI in our education

Generative AI can impact education in many ways. Currently, the technology is still in its infancy, and it is clear that schools cannot yet make a long-term estimate of the potential benefits and downsides of using generative AI in their courses. Therefore, some caution and restraint are appropriate before deciding to adopt AI tools in our education. However, these tools are available, and we must operate under the assumption that students are using them. In the short term, we see mostly a threat for assessment.

 Be aware that generative AI tools are provided by commercial entities. Any data you enter is being stored and can be used to further train the model. Therefore, you should not upload any private, sensitive, or confidential material. This also means that you cannot force students to use any of these tools. If you decide to adopt generative AI in your course, you will always have to provide an alternative to students.

Key take-aways

- Generative AI tools like ChatGPT and Bing Chat will have an impact on our education.
- Currently, the biggest impact of generative AI is on our ability to assess assignments. If your course uses assignments, please use the AI assessment guidelines within your own faculty/school.

- If you want to integrate tools like ChatGPT into your course, please reach out to your [Learning & Innovation team](#)
- [Opens external](#)
- .

What does this mean for your courses?

In the short term, we recommend that most courses do not significantly change their content or set-up in a response to generative AI. In making this recommendation, we also recognize that for some courses generative AI could be an excellent tool to improve education. For these courses, it may be beneficial to actively integrate generative AI in teaching or assessment.

Irrespective of your decision to integrate generative AI into your course, or not, some active involvement on your side is required. For any questions, please reach out to your [Learning & Innovation team](#)

[Opens external](#)

to discuss the best way forward.

Misuse

Within each faculty, the [examination board](#) has an important role in dealing with and preventing fraud. An examination board is an independent body within the faculty that focuses on monitoring the quality of education. When a student uses AI software such as ChatGPT without permission from the examiner (the person who develops a course and its exam), this is considered [plagiarism or ghostwriting](#) by the examination board. Appropriate sanctions may follow. The specific sanctions depend on the circumstances and the policy of the faculty concerned. Each faculty and each case is therefore handled differently.

Misuse of generative AI

There are three ways in which students can misuse generative AI:

1. Generative AI is not allowed, but the student still makes use of these tools. You can compare this to ‘ghostwriting’.
2. Generative AI is allowed, but the student uses it in the wrong way. For instance, the student does not explain and reflect on the use when you made that a requirement. You can

compare this to wrong references. Keep in mind that referring to generative AI is new for students. Mistakes may happen.

3. Students use generative AI for ‘traditional fraud’. In other words, the student uses generative AI to rewrite existing material (internet sources, or an assignment from their fellow student) without appropriate referencing.

When to suspect the use of generative AI tools

There are some indicators that a text may be written using generative AI, these include:

- Incorrect, fictitious references
- Overly structured, unnatural feeling text
- Internal inconsistencies in reasoning
- Factually wrong, but confidently written texts

Presence of these indicators does not prove the use of generative AI, nor does the lack of them disprove it. The most effective indicator will be to know your students and their progress.

Using Turnitin to check for the use of generative AI tools

When students submit an assignment, you can check for the use of generative AI using Turnitin. Note that the score generated by Turnitin can only be used as an indication that generative AI was used but does not provide certainty.

How does it work?

- Check the [AI writing indicator](#)
- [Opens external](#)
- . This is shown in the right-hand column of Turnitin's similarity report.
- If the AI writing indicator shows a positive percentage, click this percentage to open the AI writing report. In the report, the sections that were likely written using generative AI are highlighted. This can give you a better understanding on how students may have used generative AI.

 Note that EUR does not have agreements with other tools (such as GTPZero, OpenAI Text Classifier, etcetera). Please consult with your [Learning & Innovation team](#)

[Opens external](#)

before uploading student data to those platforms.

Limitations of Turnitin's AI Writing Detection

Step-by-step AI assessment guideline

Determine the risk of generative AI on the ability to assess your students. Based on the results, you may choose to make changes to the way you assess. If you have questions or need advice, contact your [Learning & Innovation team](#)

[Opens external](#)

.

1. Determine the risk in your assessment

☒ Exams are not affected by generative AI. These are supervised in a controlled environment, and students cannot use these tools. There is no need for further action.

☒ Other types of assignments are potentially affected by generative AI. Proceed to step 2.

2. Use generative AI to complete your assignment

Try to complete your assignment using [ChatGPT](#)

[Opens external](#)

(for writing well-structured texts) and [Bing Chat](#)

[Opens external](#)

(for information about the content).

Remember that the output you obtain is only one possible outcome. Depending on the use of these tools, other results are also possible. Really take some time to practice with generative AI to experience what it can do.

Some tips on using generative AI

- Be specific: The more specific your prompt, the more accurate the response you're likely to get. Instead of asking broad questions like "What is the meaning of life?", try asking specific questions like "What are the major philosophical theories on the meaning of life?"
- Provide context: When you provide context for your prompt, generative AI can better understand what you're asking and provide a more accurate response. For example, you could first enter some content from your course, before asking ChatGPT to complete your assignment.
- Experiment with different prompts: Try different prompts and see what kind of responses you get. Experiment with different phrasings and wordings to see how ChatGPT responds.

- Make the problem smaller: Instead of simply copying your entire assignment, ask ChatGPT to complete a specific part of your assignment
- Provide feedback: If ChatGPT doesn't provide the answer you were looking for, provide feedback on the response so that the model can learn and improve over time.
- Difference between ChatGPT and Bing Chat: Bing Chat provides more accurate answers. However, it is tuned to provide brief answers. ChatGPT is better suited for (re)writing longer texts, but lacks some of Bing's accuracy.

3. Evaluate the output in relation to your assignment

- How does the generated output relate to your learning objectives of your course?
- Does the level of the output match the intended level of your learning objectives?

Mainly focus on the content of the output produced by generative AI.

Indications

- ☒ The level of the output is insufficient, or it is unusable for your assignment.
- ☒ The output is usable and helpful but not fully at the level of the learning objective.
- ☰ The output is usable and at the level of (or exceeds) the learning objective.

4. Consider assessment criteria in relation to the output

- Is it still possible to assess the student's own contribution using the assessment criteria?
- Will you be able to differentiate between students simply using generative AI and students adding their own insights?
- Using your assessment criteria, can you still check that the student has met the learning objective?

Indications

- ☒ Using the criteria, the ChatGPT output is essentially irrelevant. You can still fully assess the student's own contribution.
- ☰ Using (all or some of) the criteria, it is not possible to assess the student's own contribution.

Assessing writing skills

If you assess writing skills, be aware that generative AI can always be used to:

- Develop an outline for the assignment and help students to structure texts.
- Paraphrase literature and other sources.
- Improve quality of writing (spelling, grammar, flow) and tone of voice.

Communicating an AI-course policy

You will need to communicate to students if and what use of generative AI is allowed for your course/assignment. You can follow the steps below to determine what use fits with your assignment and learning objectives.

1. Determine the allowed use

Before you get started, remember that some use of generative AI cannot be ruled out. Prohibiting it is unlikely to be an effective approach. Firstly, enforcing such policy for non-supervised assignments is difficult, if not impossible. Secondly, in the AI assessment guideline you mitigated the impact of generative AI.

Instead, using generative AI as support (writing, brainstorming, etc.) should be fine for most courses. Still, clearly explain what is accepted and what is not. Moreover, you could advise against (as opposed to prohibiting) using generative AI for (parts of) your assignment.

2. Explain why you decided on the policy

This step may be most vital to get the desired result. Clearly explain and discuss with students why you recommend a certain use of generative AI.

In particular if you recommend students not to use generative AI, make sure they also understand why they benefit from doing the assignment without it. What do they get out of the process?

3. Decide on the need for referencing, reporting, and/or reflection

You may want students to clearly indicate if and when they used generative AI. Some options include:

- Referencing the use of generative AI
- Report the use of generative AI in a separate section of the assignment. The student should explain how they used the tool.
- The previous option can be extended to include some reflection, for instance: why did the student use generative AI? Was the result useful?

68. Responsible use of generative Artificial Intelligence

Responsible use of generative Artificial Intelligence

KU Leuven is open to the use of generative AI (GenAI)-technology concerning education and research and encourages her students, teaching staff and researchers to handle this technology in a responsible and critical way. GenAI-tools have found their way into the university and it is of importance that everyone understands how GenAI works, to ensure that the academical standards are upheld and users maintain ownership over their written text.

What is generative AI?

Generative Artificial Intelligence (GenAI) is a type of machine learning. It's the umbrella term for a group of algorithms that can create new content. This content can take different formats: text, code, images, videos, and music, or a combination of all of these.

GenAI generates output in response to a query/prompt using generative models such as Large Language Models (LLMs), relying on large datasets to achieve this.

Some well-known examples are text generators such as ChatGPT, ChatGPT's integration into Microsoft Bing, and image generators such as DALL-E and Midjourney.

Basic principles

Transparency about the use of GenAI depending on the type of use.

Verification of the correctness of the generated output, with attention for correct sources.

Respect for copyrighted material, personal data and confidential information (including unprotected IP) by not importing them on platforms managed by external parties (non-KU Leuven servers). This is only possible with explicit approval of the owners of the copyrighted data, information or material.

Responsibility for the correct usage of GenAI (primarily as help and support) and for the output you publish (concerning research) or submit as a student

Specific principles and guidelines per target group

In addition to basic principles, KU Leuven employs indispensable principles and guidelines adapted to the target group. These are continually updated with new information and insights consistent with rapidly evolving technology. Be sure to go through them before getting started with GenAI!

Student: <https://www.kuleuven.be/english/education/student/educational-tools/generative-artificial-intelligence>

Teaching staff

<https://www.kuleuven.be/english/education/leuvenlearninglab/support/highlighted/generative-artificial-intelligence>

researcher

<https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genai/genAI>

Student Teaching staff Researcher

The necessity of correct referencing

Being transparent about the use of GenAI is essential. After all, you are influencing AI-generated content by formulating prompts, and it is not straightforward to reconstruct the information source. Therefore, with authorised use of GenAI, it is important to correctly attribute any acquired information or ideas following the guidelines set by publishers, teaching staff, or reference style guidelines.

In many cases, when using a GenAI tool, you should refer to it as ‘personal communication’ or unpublished text. It is also recommended to provide the drafting date, allowing the reader to know when the AI content was generated, or you should refer to the used GenAI model’s version.

Manual reference styles

Tips and considerations

Be sure to keep in mind some considerations of GenAI-tools while exploring its possibilities.

[Consult our tips to use the tools in a responsible way](#)

69. AI in education, resources for teaching faculty

AI in education

Resources for teaching faculty.

Arguably, Artificial Intelligence systems have been around since 1943, but now, 80 years later, have suddenly gained unprecedented prominence with the release of one particular system, ChatGPT.

Rapid releases of new systems and versions are to be expected. While there are some general ideas about Academic Integrity, it is challenging to write definitive guidelines, tips, tricks, and rules for their usage across all assignments, courses and curricula.

Instructors are strongly encouraged to set rules and provide guidelines for their courses and for individual assignments, projects, and exams within their courses (this was also frequently desired by students on survey regarding the use of AI at ETH).

Exploring the use of AI in teaching and learning is an active field of research and development. The rectorate is involved in a number of projects, and ETH-internal funding for new projects can be obtained through the Innovedum Focal Point Theme AI in Teaching and Learning. We are also collecting additional resources.

The revised DownloadDeclaration of Originality (PDF, 183 KB)vertical_align_bottom (DownloadEigenständigkeitserklärung (PDF, 175 KB)vertical_align_bottom) is now available.

Course at ETH Library: [Scientific writing – Using ChatGPT effectively and responsibly](#)

ChatGPT: <https://ethz.ch/en/the-eth-zurich/education/ai-in-education/chatgpt.html>

Projects <https://ethz.ch/en/the-eth-zurich/education/ai-in-education/projects.html>

Academic Integrity <https://ethz.ch/en/the-eth-zurich/education/ai-in-education/AcacIntegrity.html>

Resources <https://ethz.ch/en/the-eth-zurich/education/ai-in-education/resources.html>

70. AI tools and your studies

At the UvA, students and lecturers are not (yet) allowed to make active use of AI tools for teaching and assessment. However, the UvA is exploring how it could use Artificial Intelligence as a teaching aid in the future.

About ChatGPT

ChatGPT is a language model that has been trained to generate text rapidly and answer questions based on user prompts. Users can ask follow-up questions or give further instructions to better tailor its output to their requirements.

Why does the UvA not use ChatGPT for teaching purposes?

ChatGPT is a commercial product. It's currently unclear what the developer does with user data and any other data entered into it. This means that the tool is not (yet) suitable for use at the UvA. For this reason, lecturers can't require you to use ChatGPT to complete assignments.

AI tools may not be used to write assignments you'll be submitting

As an UvA student, you can expect to receive high-quality and innovative education. At the same time, you must have the intrinsic motivation to learn. You'll still need to write your assignments yourself, instead of letting an AI tool take care of this. This helps develop the skills you'll need later on in your studies and on the labour market.

Assignments that you didn't write yourself may be deemed fraudulent

The [UvA's fraud and plagiarism regulations](#) state that lecturers must be able to assess a student's knowledge, insight and skills. We therefore expect students to write everything they submit

themselves. Unless explicitly stated otherwise, the use of AI tools such as ChatGPT is therefore not permitted. Consequently, submitting assignments that you didn't write yourself may be deemed fraudulent. In cases of fraud, the UvA will take strict action.

Risks when using AI tools

AI tools can help you to study more effectively. For example, you can use ChatGPT to brainstorm, to check your knowledge when studying for an examination or to translate a text for you. It can be a convenient assistant that is available at all times. However, it's important that you're aware of the risks involved.

- Answers that sound perfectly plausible may still contain inaccuracies. ChatGPT generates output based on probabilities and statistics. It doesn't verify whether the information is factually correct.
- The output may be compromised by harmful bias and stereotypes. Among other things, this is because the data sets that ChatGPT uses are not representative.
- Many AI tools store your interactions, which entails privacy and intellectual property risks. You should therefore take care not to enter any privacy-sensitive or other confidential information, such as confidential research data, patient information or personal data of fellow students or lecturers. Not even a paid account offers adequate security and privacy safeguards.
- Don't use your UvA account, but use a dummy email address that can't be traced back to the UvA, and ideally can't be traced back to you personally either.

had developed an e-learning module for students in English about the responsible use of AI tools (particularly ChatGPT) in higher education. This interactive e-learning module consists of text, knowledge clips and brief knowledge quizzes. It will take around 45-60 minutes to complete.

[Go to the e-learning module](#)

71. Guidelines to use artificial intelligence at UiO

The development in artificial intelligence will impact our learning, knowledge development and work methods. On this page, we will communicate perspectives, principles, and insights on the use of AI in higher education.

Approved AI tools at UiO: GPT UIO: <https://www.uio.no/english/services/it/ai/gpt-ui/o/index.html>,
Whisper:<https://www.uio.no/english/services/it/research/sensitive-data/help/hpc/software/whisper.html>,
Autotekst: <https://www.uio.no/english/services/it/video-sound/autotekst/index.html>

How to use AI as a teacher

On this page you will find advice, considerations and approaches to learning and assessment with the use of AI, primarily large language models, as teachers.

Understanding AI and the possibilities the tools provide

When we use the term "AI" on this page, we primarily mean artificial intelligence based on large language models, such as OpenAI's ChatGPT and similar tools. If you want to use AI tools for teaching purposes, you should use GPT UIO for privacy reasons. In order to use AI in a fruitful and responsible way, it is crucial that you familiarize yourself with the possibilities and limitations such tools provide. This means that you should have familiarized yourself with [ethics and social responsibility](#), as well as [legal perspectives](#) before you or your students use the tool. The students also have their own [information page about AI](#). There, the students are encouraged, among other things, to inquire with you about acceptable AI use in your subject. In "recommended resources", you will find selected material that can help you further understand AI. Here you will also find specific websites you can follow to stay up to date.

What are language models

Generative Pre-trained Transformer, abbreviated GPT, is a type of artificial intelligence model developed by OpenAI to generate text. The GPT model is trained on large amounts of text data and learns patterns and language so that it can generate coherent and relevant text based on what you write to it. You can read more about what language models are on the website "[What is GPT UIO](#)".

It is important to be aware that the language models are not knowledge databases. Nevertheless, with good instructions, the tool can be used in various ways linked to many domains of knowledge. However, it is crucial not to take for granted that the answer you receive is true or of sufficiently good professional quality.

Examples of AI tools

There are a number of tools that use different forms of AI to perform tasks other than what GPT UIO can do today. You should therefore be aware that students potentially use more AI tools than those offered by UiO. An example is the plus subscription to ChatGPT which gives the opportunity to both analyze and create images, do web searches, create small programs and run them to solve mathematical questions or use third-party tools to solve other tasks, such as for example looking up databases, connect to web services or perform other specialized tasks with or without AI. In addition, there are specialized AI tools in a number of fields such as video, music, images, and more. See the page [Futuretools](#) for a comprehensive overview.

Explore the use of GPT UIO

Once you have understood the basics of the technology, the next step is to understand, test and experience how the tool can be used in your discipline. It is crucial that you put yourself in the driver's seat and have a clear idea of what you want to get out of the tool.

If, for example, you want to use AI for teaching planning, relevant information such as the teaching's aims, content, framework and participants can help to give more precise answers than if you omit this information. In many contexts, it will also be important to share information about forms of assessment and other activities, as well as your experiences about what is important and difficult for students to understand and master in the subject.

Prompting

When you talk to language models, it is important to be able to ask concrete questions and give tasks that contain sufficient context. This action is called prompting. Prompting is your way of giving GPT UiO instructions about what you want it to do. It can be challenging to create prompts that give you the desired response. GPT UiO bases a lot of the answers it gives on the information and context you give it. There may therefore often be a need for adjustments, corrections and follow-up questions during the conversation. This means that the better your understanding of the topic, the greater the probability that you can create useful prompts. In order for such AI tools to be perceived as useful for you, it is important to view the interaction as a dialogue. Such tools are therefore not intended to be used as a search engine. At the bottom of "recommended resources", you will find selected material that can help you understand prompting.

Tips for creating good prompts

1. **Give GPT UiO a role** - for example, instruct the model to behave like an expert in the field you want help with.
2. **Be specific** - consider giving one instruction at a time to get more precise answers, than if you give it many tasks in the same question/task formulation.
3. **Rephrase** - if you feel that the answer GPT UiO does not give correct answers or is on the side of what you are asking; try to rephrase the question by adding more details.
4. **Use examples** - to sharpen the answers GPT UiO gives, consider exemplifying what you want an answer to.
5. **Context** - define the context of the conversation. For example, you can ask GPT UiO to create a case assignment based on a given topic. You must then make sure to explain at which level the case is to be solved - eg BA/MA, the teaching context, how long the students will spend, length of the case and which theories you think should form the basis of a proposed solution.
6. **Think step by step** - language models generate the most precise answers if you instruct it to explain step by step.
7. **Ask for help in asking questions** - you can ask GPT UiO for advice on how to ask more precise questions. In many cases, it will ask for more details related to the question you have asked, which can help you get a more precise response.

- 8. Get tips for areas of use** - you can also start by writing the job title and your tasks, and then ask what GPT UIO can help you with. You will then receive a list of suggestions for possible areas of use that you can use as a starting point for further conversation.
- 9. Test again in a new thread** - When you test the same prompt in a new thread, you may find that the answers vary, since AI can generate different perspectives each time. That is, the "memory" of GPT UIO is limited to the individual conversation thread you are in. By using multiple conversation threads, you can explore different response styles and points of view, and the responses can be perceived as more relevant. In addition, one can avoid potential confusion from previous context, if the thread of conversation was long.

Assessment activities

AI re-actualizes the importance of developing abilities such as critical reflection, independent thinking, creativity, assessment skills, nuance, as well as the ability to wonder and ask questions. These have always been important goals in higher education. There may be good reasons for allowing the use of AI in student work to be assessed. At the same time, it can represent a challenge to assess what the students have actually learned. This means that we have to discuss which forms of assessment are suitable for bringing out such abilities in your subjects. What skills and knowledge are you going to test the students on? How do you know if the students have learned what they are supposed to?

Exam answers and cheating?

There are no tools that can reliably identify AI-generated text in an answer. Although it is claimed that some tools can, AI is developing so quickly that such a "plagiarism check" is continuously out of date, and therefore cannot be considered reliable. In addition, there are always ways to fool the plagiarism check.

Our recommendation is therefore to focus on developing forms of assessment where the main focus is on the process of learning, as opposed to only focusing on the final product. Such forms of assessment are important for the entire course design, and are in line with what the research highlights as strengthening for learning. [UiO's regulations](#) allow for many different forms of assessment, so there are options. Below we have listed some suggestions, and you can find inspiration in the section "learning activities with AI?". In some cases, there may be a need for a

school examination, either during a transition period, or as a continuous form of examination based on the learning objectives that have been chosen.

UiO's examination regulations focus on the student's answer being their own work. With AI, this compliance principle is challenged, and clear guidelines at the faculty are desired. There are many subject-specific assessments that must be made. For more perspectives, the [AI ethics page](#) is recommended.

Under "recommended resources", you will find selected resources that can support you in the development of assessment and learning activities.

If you allow the use of AI

Provide clear guidelines for the use of AI, and emphasize the importance of taking privacy, copyright and ethics into account

Clearly communicate the requirements for academic integrity, accuracy and use of sources to the students

Clearly communicate if the use of AI is expected or only a possibility

Give students insight into risks of using AI

Require transparency. That is, description and documentation of the process. See what we write about this on the AI student page.

Require reflection on the process and use of the tool

Require use of adequate sources

Give the students training in the use of the tool, either as physical training or reference to relevant user documentation.

Clear examiner guidance that explicitly addresses the use of AI

Proposals for assessment activities

When designing assessment activities, it is important that the tasks the students are given cannot simply be solved by AI alone. Draw on experiences and reflections that students do or have done into the assessment activities you design. As mentioned earlier, in several cases it will be beneficial if the students also visualize how they have worked with the assignments, by describing their reflections during the process and what they have learned.

Connect tasks to concrete work or experiences

Connect the task to concrete work in teaching. For example, you can work on one or more cases in class, and questions can be linked to this specific work. The student work with cases can be included in the assessment basis, for example through a portfolio assessment.

Connect assignments to student field practice experiences.

Connect to experiences. Let the students explain and reflect on process execution, learning activities and problem solving.

Assignments that ask for the student's own views. Give tasks that involve analyzing complex issues, assessing alternative solutions and arguing their own position.

Assignment formulations linked to the syllabus. Create complex task formulations specifically linked to the syllabus, which require critical reflection. For example, students can be asked to use the syllabus to highlight actual/practical/current issues.

Assess and/or compare different texts using the syllabus. This can, for example, be the work of fellow students, fictitious student answers, published articles or texts produced by GPT UiO.

Have the students create something new. Give tasks without predefined answers and encourage originality and creativity. For example, students can be tasked with developing and justifying new research questions or justifying arguments, with references to specific literature.

Iterative forms of examination

These are examination forms that are divided into several steps over a longer period of time, with feedback rounds and opportunities for the student to improve their own work.

Mutual assessment. Fellow students give and receive feedback on each other's work along the way and improve their work before final delivery. During the process, AI can contribute feedback as a "fellow student".

Folder assessment. The students must work in several steps before delivering an academic final product. AI can or should be used as a supervisor in the process.

Project work. Parts of the work take place in class, and the students receive feedback along the way to improve their work. This work may contain submissions/presentations along the way.

Logs and reflection notes. To make the students' thought processes visible, they can submit logs or reflection notes in which they explain their use of GPT UiO. Here it is important to bring out the reflections and assessments they have made at each step. Examples of questions can be

How did you use AI to arrive at the result?

What kind of prompts/instructions did you use?

Which prior knowledge and skills did you notice were important and useful in the process?

What have you learned from the process? What experiences do you take with you from the chat you had with GPT UiO, and why?

Multimodal forms: More than words

Oral forms. For example, video reflection, presentation and/or conversations.

Multimodal responses. As part of the answer, students can develop content in formats other than text, for example video reflections, video essays, podcasts, presentations, interviews, visual representations.

Multimodal tasks. Create tasks that consist of different formats in addition to text. However, Chat GPT-4 is capable of interpreting and producing some visual material. Development is rapid; as of now GPT UiO cannot do this, but that may change in the future.

Video case exam. The students watch a video that thematizes a topic from the syllabus. The students are asked to analyze the video with the help of the syllabus.

Learning activities with AI?

In higher education, reading original sources, writing one's own texts and independent exploration and problem-solving are essential in order to be initiated into one's subject tradition and knowledge practices, learning, reflection and critical thinking. There may therefore be good

reasons to include such activities in a way that is fully or partially shielded from AI, but there may also be good reasons to include AI in such activities.

The connections between learning outcomes, assessment forms and learning activities should be communicated in a way that enables the students to think about their own learning. Such awareness of one's own thinking and learning is called metacognition. If you want to use AI tools, you should include AI in such a way that it strengthens the students' learning and metacognition, and at the same time their understanding of the technology and how it can be used appropriately in the learning process. We recommend that the learning activities are designed as an iterative process, where the students document any dialogue with AI, and justify the steps and assessments they make. In addition, the students should work together so that they can discuss and practice evaluating the AI-generated content.

At the bottom of "recommended resources", you will find selected material that can support you in the development of assessment and learning activities.

What roles can AI have in teaching

In line with the development of AI, new areas of use are constantly giving new areas of use you can consider using in your work with teaching. Important mental work is of course something you can't dismiss when it comes to planning activities and teaching, formulating assignments and the like. You as a teacher know what the students will learn. Therefore, it is crucial that you are in the driver's seat and have clear ideas about what the students will learn from your teaching. What are the central issues, concepts and sources in the subject? What prior knowledge do the students have, and what might be challenging for them to understand and master in the subject? You may also include this in your instructions to GPT UiO. In this way, cooperation with GPT UiO can be awareness-raising for you as a teacher. Remember, it is you as a teacher who are responsible for ensuring the quality of the AI generated content.

Here are some examples of tasks you can use as a starting point to explore the GPT UiO

Development and testing of assignments and cases

Development of multiple-choice questions or other quiz formats

Development of learning activities

Development of presentations and scripts

Translations

Suggestions for learning activities without AI

When you are with your students in the classroom, you can set conditions for how the learning activities should be carried out, and whether they should be done with or without KI. When the students are at home, you do not have the opportunity to control this. Therefore, you should consider giving students assignments and learning activities, AI cannot easily solve. In both cases, it is crucial to communicate to the students the purpose and how the activities are intended to contribute to their learning. This kind of metacommunication about your expectations helps students understand your learning intentions. This benefits their metacognitive abilities.

Activities in the classroom

Writing activities

Peer assessment

Reading activities

Discussion activities

Problem solving

Case assignments

Lab activities

Activities both inside and outside the classroom

These are activities where AI cannot easily generate the central parts of the answer

Tasks which are complex and specifically connected to the students' own experiences and reflections, for example from an excursion, a seminar, a lab exercise,

Step-by-step assignment with feedback from fellow students and discussions along the way

You can find more examples and inspiration for learning activities in the university pedagogical online resource [BetterTeaching](#), see especially the modules "Teaching planning" and "Assessment and feedback".

Proposals for learning activities with AI

GPT UiO can be used as a learning partner as a way for the students to have an ongoing, scaffolding, academic dialogue with the tool. The purpose of such learning activities is for the students, and preferably together with fellow students, to build on or create ideas together with GPT UiO, or for GPT UiO to ask questions in such a way that the students have to reflect on outstanding knowledge and learning processes.

Such learning activities can be

Devil's advocate: GPT UiO is asked for counterarguments to a proposed solution or a series of arguments.

Idea generator: GPT UiO can be used as a partner to generate ideas, for example for projects or issues for discussion.

Questioner: GPT UiO is asked to give the student one question at a time related to a subject they want to be tested on - like a live test/quiz. GPT can give feedback on the answers the students give.

Feedback task: Group assignment with GPT UiO where the students will receive feedback from GPT UiO on subject texts they have developed, or other texts they upload. The students discuss improvement potentials in the texts, preferably based on assessment criteria and/or examiner guidelines.

Critical assessment of GPT-generated content: Content generated with the use of GPT UiO must always be quality-reviewed using credible sources. It can also be useful for students to compare GPT-generated text with other subject texts. Students learn by comparing the academic basis against the AI-generated content.

Co-producer in creating learning activities: As part of the teaching, the student can be tasked with creating learning activities together with GPT UiO and fellow students. For example, case assignments, discussion assignments, flashcards, multiple choice questions, board games etc.

On the [AI student page](#) you will find examples of how students themselves can use KI as learning support.

How to Use AI as a Student

On this page, you will find examples and advice on how you as a student can use AI, in ways that promote learning.

Learning with GPT UiO

GPT UIO is a useful tool that can contribute to your learning outcomes. However, if you let the tool do the work for you, you could miss out on a lot of valuable learning. It is you as a student who should manage your own knowledge development.

Excessive use of AI-generated texts can hinder your development of analytical skills and critical-and creative thinking. Writing and editing one's own text is often important for developing subject understanding and one's own thinking in addition to your academic writing skills. The purpose of writing assignments at the university is not primarily to assess what you have learned, but to give you opportunities to develop your skills in academic writing and thinking. It is incredibly important that if using AI, you put yourself in the driver's seat of the writing and use AI as a tool and a support to become better at writing.

GPT UiO can create summaries and simplifications of large and difficult texts. This can be useful. But it is also very important to read the texts yourself and process the content. This is how you get a good understanding of academic work in your course and can evaluate the content and what it means. When working with text-generated content, it may be wise to keep these questions in mind

Can I trust the answers GPT UiO gives?

Do I have the opportunity to check the content against original texts?

Should I ask the question in a different way to get a more extended or specific answer?

Understand the Technology

It is particularly important to understand that the technology behind some AI tools, such as GPT, is based on language models. This means that they are not knowledge-based, but the answers they provide are based on statistical forecasts that the model is trained on. The language model

looks at patterns in the words you use in your questions and calculates what the model thinks are the most important words in the sentence to give an answer that is likely to make sense. For example, the model can calculate whether the word "bark" in the term "the bark was old and gnarly" means the protective outer covering of a tree or the sound a dog makes.

compares this calculation with the statistical pattern the model has from training.

calculates the probability that certain words may be appropriate in an answer and gives you the most likely words back.

Remember that if some words are likely, does not mean that the answer is true. If you ask a language model to complete the sentence "The Earth is...", and the language model is trained on many texts containing the sentence "The Earth is round", it will suggest the word "round". If it is trained on texts that say the Earth is flat, it will suggest that instead.

No Definite Answers

In other words, there is no form of fact-checking involved when a language model like GPT UiO responds to your questions, just a probability calculation based on texts it has analyzed. We do not know which texts GPT UiO has been trained on, neither do we know exactly how it has been trained. The same applies to most other language models. Therefore, we cannot rely on the information they provide us as true.

Basics and Useful Resources

Here are some entrances to a basic understanding of the technology. Once you have understood the basics, the next step is to understand and experience how available tools can be used in your course. See our selection of resources that can help

1. The video [Introduction to AI for Teachers and Students](#). Produced by Wharton School.
The video explains what artificial intelligence is. The video illustrates other AI tools than GPT UiO, but is based on the same language technology. Viewing time is 10 minutes. Available in English.
2. The website [AI ethics and Social Responsibility](#) presents the critical aspects that you as a student at UiO need to consider. Estimated reading time is 10 minutes. Available in Norwegian and English.

3. The video [Prompting AI](#). Produced by Wharton School. The video shows how you can ask questions, read prompt, in a useful way. The video shows other AI tools than GPT UiO, but you can use the same approach. Viewing time is 11 minutes. Available in English.

Learn more

The video [Large Language Models](#). Produced by Wharton School. The video gives you an explanation of large language models. The video illustrates other AI tools than GPT UiO, but is based on the same language technology. Viewing time is 12 minutes. Available in English.

The report [ChatGPT and Artificial Intelligence in higher education: Quick start guide](#) issued by UNESCO gives you a short description of what artificial intelligence is, examples of use, and outlines some critical reflections one should be aware of. Estimated reading time is 30 minutes. Available in English.

The website [What are AI tools and how do they work?](#) created by University of Groningen gives you a short explanation on how AI tools function, and which AI models exist. Estimated reading time is 5 minutes. Available in English.

Explore and Use

Before you start, ask your teachers what they have decided is acceptable use of AI in your course. Then, explore with fellow students how this technology can be used to support your work. Here are some questions that might be relevant to work with in such a process

How can the tools support learning in your course?

What are the tools good at? What are they bad at?

How can using the tools contribute to, and elevate your learning?

What is important so that the use of the tools does not make you learn less?

What ethical concerns are relevant?

What does ethical concerns mean for how you use the tools?

Prompting

When speaking with language models, such as GPT UiO, it is important to be able to ask specific questions that contain a lot of context. This action is called prompting. In Norwegian, prompting can be explained as question formulations, instructions, or cues. Prompting is your way of giving GPT UiO instructions on what you want it to do.

It can be challenging to create prompts that provide you with the desired response. GPT UiO bases much of its answers on the information and questions you provide. Therefore, there is often need for adjustments, corrections, and follow-up questions in your prompts. The stronger your grasp on the subject matter, the higher the likelihood of producing useful prompts. The most important thing is not to ask the perfect questions, but to try out and test. In some cases, you might get more accurate results if you prompt in English. For such AI tools to be useful to you, it is important to view the interaction as a dialogue. If you want to learn more about this, go to the "Good to know section" on this webpage.

Proper, critical and honest use

Make sure you are well acquainted with the guidelines in your course. Ask if you are unsure. In general, this applies to all work that is assessed. For example, exams should be your own work, even though you are often allowed to use exam support materials. It is your teacher who decides if and how you might use AI.

To use AI-generated text in an academically responsible manner in your writing process, you must actively adhere to academic norms for honesty, accuracy, and transparency. This means that it is important that you

- clearly indicate when and how you use such texts
- critically assess and verify information from AI sources
- do not rely on AI as the only or primary source of information

If the task consists of writing a text, this will, for example, entail that you show in your text which sources you are building upon. You must further clarify for the reader the difference between your own voice and the voices of the authors you are referring to. You can read more in [search and write resource](#) on how you can practice clearly expressing your own opinions and what you are building upon from others.

If you use text generated by AI, for example GPT UiO, in the process of developing your own text, you must be aware that such AI-generated texts do not adhere to academic norms for accuracy and transparency.

AI-generated text may contain errors, inaccuracies or be misleading. So always verify the text with several other sources.

AI-generated text is not your own. If you use it in your work, you need to be open about which parts of the text is AI-generated, how it was generated and used in your work.

AI-generated text usually does not refer to sources, or sources it refers to are not necessarily real or relevant. . To be an honest writer, you need to find, explore and reference real academic sources.

AI-generated text may reflect biases or prejudices from the training data. By building on biases or prejudices ,you may contribute to reinforce such bias and prejudices. .

Do not use AI to write the text for you, but as a support in the writing process, for example to get ideas and improve your own text.

If you use AI in the process of developing your own text, you should be open about this, and clearly show how you have used AI and what is your independent contribution in the development of the text. Here you can [read more about how you can be transparent](#).

Principles for the Use of Artificial Intelligence (AI) in Education

Whether or not you incorporate AI into your teaching, it is essential to understand certain principles and consider their significance for your specific teaching approach and for your students.

1. Safe and Responsible Use of AI

Familiarise yourself with UiO's guidelines for data protection and ethical use of AI tools before using, for example, GPT UiO. It is also your responsibility to inform students about these guidelines so that they use the tools responsibly.

2. Choose AI Tools that Support Students' Learning Outcomes

Reflect and make conscious decisions about how the AI tool supports the learning objectives and learning activities for your subject. Communicate the purpose to the students so they understand

what they are to learn. You should spend time getting to know the students' level of knowledge, experiences, and technological understanding before you implement artificial intelligence.

3. Equal Access

Students will largely use AI tools in or outside of teaching. Therefore, all students should have equal access to necessary information about tools to be used in teaching and guidelines for proper use. Offer training or link to documentation so that everyone has the same prerequisites for use.

Additionally, some AI tools will serve as support for students with special needs. Provide extra guidance and training if students are expected to use a tool they have little or no competence in. Allocate time for testing, collaboration, and feedback sessions as part of the teaching, where you and the students can support and help each other.

4. Promote Critical Thinking

Encourage students to reflect on and gain a critical understanding of AI use, including limitations, opportunities, and ethical dilemmas. Being able to critically assess the material presented by an AI tool is, and will be, a necessary skill in today's and tomorrow's working life.

5. Speak with your students

Having an ongoing dialogue with students about how they have used AI tools will give you valuable insight into their knowledge and pre-understanding. At the same time, students will benefit from critically reflecting on their learning methods.

6. Stay Updated

Stay informed about new AI tools and pedagogical approaches that can improve your teaching.

Collect feedback and evaluate students' experiences with AI in teaching. You can use their experiences to adjust learning activities along the way and to plan future teaching.

Talk to colleagues. Ask for advice on AI activities you have planned. Perhaps you also have experiences that your colleagues can benefit from?

Be proactive. Familiarise yourself with AI technologies by trial and error. Learn how the tools work by testing different questions (prompts), both through the teacher and student roles.

Keep an eye on what's happening at UiO regarding AI. Both locally at your faculty/institute and centrally, various forms of training related to AI in teaching will be offered during the fall semester. Stay updated on ui.o.ai.no.

Ethics and Social Responsibility

When using artificial intelligence-based technology, you must make some ethical reflections. Here you will find examples of questions you should ask yourself before using the technology, while using it, and most importantly, when assessing the outcome of its use.

Artificial intelligence offers many exciting opportunities in studies, research, and other work. However, there are many pitfalls you might fall into if you use the technology uncritically. You must reflect on what is good and bad or acceptable and unacceptable use. In other words, you need to make some ethical considerations.

Important questions you should reflect on

Can I input personal information into an AI tool?

Data protection requires that you have good control over the personally identifiable data you process in research or studies at University of Oslo (UiO). You cannot input personal information unless it is permitted by law, UiO has clear [guidelines on this](#).

Do I know who is behind the AI tool, and what interests are at play?

Are the owners, or the sender, clear about how the AI tool works, how it has been trained, and how it uses the data that is inputted? If the AI tool is transparent in how it was built and processes data, it's easier to assess whether it can be used in studies and research.

Can I input someone else's text or results into the AI tool?

Other people's work may be copyrighted, and you cannot use this work as you please. If the AI tool uses the material inputted for further training of the AI tool, you cannot use the material in the AI tool.

Can I use AI tools to write my academic text and claim the output from the it as my own work?

If your subject allows it, you can use text-generating AI tools to develop ideas conceived by you, but the results should not replace your own work. If you have used a text-generating AI tool, be open about how you have used the tool. This might include mentioning the AI tool and how you used it, or providing a link to the conversation with the AI tool (chat). See a [recommendation from APA on search and write](#).

Have you accounted for rules and guidelines on cheating and dishonesty?

Academic ethical principles require that work that is published, submitted, or to be assessed, is a result of your own intellectual effort. Artificial intelligence can assist you in formulating and drafting text, and there will be a limit to how much help you've received before you can no longer claim it as your own text. If you present intellectual work done by others, be it another author or a tool, as your own, you can be accused of plagiarism or cheating.

When writing academic text, can you be held accountable for the text you have written and the sources it is based on? Can others replicate the results you've produced using artificial intelligence?

Does UiO have rules and guidelines about cheating and dishonesty that I need to be aware of?

[UiO's cheating rules](#) state that "When you write academic texts, you should display which thoughts and reflections are your own and which you have borrowed from others' work. This allows the reader to look up the sources you've used, verify facts, and replicate your results." Fabrication of data/research data is academically dishonest.

Can I trust the answers I get from artificial intelligence-based AI tools?

Text-generating tools like GPT UiO do not provide a reliable source base and cannot search for sources. Evaluate the answers you receive critically and verify the information and claims by checking credible sources. Additionally, you should always look up your own sources and list/reference the sources you've used.

The answers you get from using artificial intelligence-based tools might be influenced by systematic biases in the data they are trained on. To avoid such bias, you should scrutinise the answers from the tool so that perspectives are not based on biases arising from, for instance, incomplete data on gender, minorities, or demographic factors.

Text-generating AI tools, for example, might produce immoral or objectionable content.

Legal guidelines for the use of Artificial Intelligence (AI)

This is a short introduction to the aspects you will need to consider regarding data security, copyright and data privacy before using AI-based technology.

The GDPR (General Data Protection Regulation, or "[personvernforordningen](#)" in the Norwegian law) is the legal framework for data privacy in Norway. This law was passed by the EU, and governs whether and how personal information may be stored and processed.

Data Security

When using AI-based tools, you need to be mindful of which data and information you allow the tool to access. The vast majority of these tools need more resources than you have available on your own computer. In order to analyze the data you give to the AI-based tool, it will likely use servers, which are other computers that your computer is communicating with.

In addition to this, the information will be sent through the internet from your computer to the server, and new information will be transmitted back to your computer. If you want to use an AI-based tool as a student or employee at UiO, you will need to consider whether your planned use is safe and allowed by the applicable laws and regulations. This applies both to the server you plan to use, as well as the method of data transmission to that server.

If you are not sufficiently careful when considering the safety and legality of the AI tool, you may risk that the data you use with the AI tool falls into the wrong hands, or is misused.

At UiO, we have two categories of AI tools

Third-party tools run on servers that do not belong to UiO, but are run by companies that have entered a data processing agreement with us. The data processing agreement means that we can exchange information with the server safely, but it has some limitations as to which types of data we can use.

UiO's ChatGPT service ([GPT UiO](#)) is one such example of a third-party tool. The service runs on servers in Ireland. Because the servers are within the EU/EEA, and because we have a data processing agreement with Microsoft, we can safely use this service for general personal data such as names and phone numbers – in other words, green and yellow category data.

This means that you will need to be careful not to use special categories of personal data, often referred to as sensitive personal data. Special categories of personal data, for example data about a person's health, religion, or political views. Non-sensitive data, such as student names, e-mail addresses and the students' exam submissions, may be used in the GPT UiO tool.

Internal tools

Internal tools are tools that run on UiO's own servers. When the tools run on our own servers, we have more control over the data being sent to the tool. This means that we can use it with more sensitive data than the case is with external tools.

UiO's [Autotekst](#) service is an example of an internal tool. Autotekst can transcribe audio recordings to text with a high rate of accuracy, as well as translate speech to text in other languages. Because Autotekst runs on servers that UiO owns and controls, the service can be used for recordings that contain sensitive personal data. The service does not send any data to third parties, since all processing happens on our servers on campus.

However, you still need to be careful not to use unsecured networks when you send data to UiOs servers. If you are using an open network (internet access that does not require a username and password), others could be able to access the data you are uploading through the same network. More information about the different types of data, and how you can process and store each type, can be found in [the UiO data classification framework](#).

More information about which data you can safely store can be found in the [UiO data storage guide](#).

Copyright

In addition to taking care that no personal data ends up in the wrong hands, you will need to be equally mindful that you are not violating the copyright of any data you use with AI tools. If you have access to a research article which is normally behind a paywall, you are not allowed to make this accessible to others, since it is subject to copyright and not intended to be shared. If you give an AI tool access to such an article, this may allow others to access the article.

Additionally, the knowledge and/or results in the article may be used to further train the AI model, and thereby end up being included in the tool and accessible to anyone else who uses the tool. If you are not being careful, you may risk violating copyright laws and agreements in this way. This can have both legal and economic consequences. However, if you are using the UiO GPT tool, this is not a problem, as the data will never be used to further train the model.

It is also important to remember that AI tools have been trained on vast quantities of data in order to become good at their purpose. In many cases, it is not known exactly what kind of data has been used for training, and there may be issues surrounding whether or not the use of data to train the model itself was legal. For instance, this is why ChatGPT was blocked by the Italian data protection authority, as they stated that the company behind ChatGPT did not have the legal right to use the data they used for training the model.

72. Using AI to support learning

What are AI-based applications like?

ChatGPT, Google Bard, DeepL and other familiar artificial intelligence (AI) applications available online are based on large language models. More recently, language models have evolved to the point where they can produce human-like text and conversations.

They can also correct and transform text at such a high level that it can be difficult to distinguish the final result from human-generated text. In the future, more such models are sure to emerge and their functionalities will continue to evolve. It is therefore important that we take their existence into account in teaching and research.

University encourages the use of AI

The existence of large language models should be seen as an opportunity. The University encourages degree programmes and teachers to use AI in their teaching. This way, we can prepare you for a society of the future where AI methods will be widely used.

As AI brings new possibilities for producing text whose origin and reliability is unclear, it's important to use them in a controlled way. The teacher can, for example, restrict the use of AI in situations where using it would not promote your learning.

If you are unsure of whether you can use AI to support completing a task, you can always ask the teacher before you get to work.

Artificial intelligence guidelines in a nutshell

What is cheating and plagiarism?

Large language models can, as a rule, be used in teaching and as a support for writing. The teacher for the course has the final call on the topic. If there's a risk that the use of large language models impedes achieving the set learning objectives, the teacher can prohibit the use of AI (independent work included).

If you use a language model to produce the work you are returning, you must report in writing which model (e.g. ChatGPT, DeepL) you have used and in what way. This also applies to theses. Please note that you should never name AI as the author of the text or other written output. AI cannot take responsibility for the content of the text – this responsibility always lies with humans.

Use of language models is never allowed in maturity tests.

Your home faculty, degree programme, or the University Language Centre can make additional guidelines on the use of AI in their teaching.

The responsible teacher should tell the student about the principles, disadvantages and benefits of using language models. If use of AI is prohibited on the course, the teacher should explain and motivate the limits of prohibited use in writing.

Equality is a core value when planning education: ChatGPT and other large language models are not always available or there may be a charge for their use. You should never be required to use a language model that is not available for free.

If you use a large language model in a course, part of a course or examination where it is prohibited in advance, please note that this constitutes cheating and will be treated in the same way as other cases of cheating. The same rules if you fail to report the use of a language model as instructed.

When using AI, always try to be precise and follow your teacher's instructions!

In addition to the University guidelines, the use of AI in teaching and learning is governed by the ethical guidelines set by the European Commission. The guidelines, available on the EU Publication Office website, are aimed especially for teachers, but taking a look at them can also be useful to students.

What are the guidelines based on?

[The Guidelines for the use of AI in teaching at the University of Helsinki \(pdf\)](#) were confirmed by the Academic Affairs Council on February 16, 2023. Please note that they may be further specified in the light of future regulation and technological developments.

Guidelines for the use of AI in teaching at the University of Helsinki Academic Affairs Council 16.2.2023 Large artificial intelligence (AI)-based language models such as Chat GPT, Google Bard, and DeepL have evolved to the point where they can produce human-like text and conversations and correct and transform text at such a high level that it can be difficult to distinguish the result from humangenerated text. It is foreseeable that more such models will emerge, and their functionalities will continue to evolve, so their existence should be taken into account in university teaching and research. The existence of large language models should be seen as an opportunity. Degree programmes and teachers are encouraged to use AI in their teaching and to prepare students for a society of the future where AI methods will be widely used. As AI brings new possibilities for producing text whose origin and reliability is unclear, they should be

used in a controlled way. Use may be restricted in teaching in situations where the use would not promote student learning. At EU level, an AI regulation is under preparation, which will also apply to AI systems in education. In addition, there is an ethical policy on AI and its use, as well as an ethical code for teachers¹. The University's guidelines may be further specified in the light of future regulation and technological developments. University policy on the use of AI and large language models in teaching and learning 1. Large language models may be used in teaching and as a support for writing. The teacher responsible for the course decides on the use. If a risk is identified that the use of large language models in a way that impedes the achievement of the learning objectives, the use of large language models may be forbidden, including for independent work. 2. If a language model has been used to help produce the work to be returned, the student must indicate in writing which model has been used and in what way. This also applies to theses. Large language models or other AIs must not be named as authors of the text or other written output, as the AI cannot take responsibility for the content of the text. The responsibility for the linguistic and factual correctness of all written material lies with humans. 3. The use of language models is not allowed in maturity tests. 4. The Faculty council, the Degree programme's steering group or the Language Centre may make additional guidelines on the use of AI. 5. The responsible teacher must explain and motivate to the students the principles, disadvantages and benefits of using language models, as well as the limits of possible prohibited use in writing. 6. When planning your education, it is important to bear in mind that large language models such as Chat GPT may not always be available or there may be a charge for their use. Students cannot be expected to use a model that is not available for free, as this would put students at an unequal position. 7. If a student uses a large language model in a course, part of a course or examination where it is prohibited in advance or fails to report the use of a language model as instructed, this constitutes cheating and will be treated in the same way as other cases of cheating.

73. Research in Artificial intelligence

Artificial Intelligence Artificial intelligence (AI) refers to systems that display intelligent behaviour. Typically, these systems analyse input data and make decisions or take actions. The application of AI to real world devices is quickly transforming our industry, our society and our world. Specific fields: Natural Language Processing: development of models and algorithms for the analysis of written texts, with end-user applications such as text categorization, question

answering, sentiment analysis and machine translation. Preference Reasoning in Decision Support Systems: This area focuses on compact preference representation formalisms, preference aggregation methods in multi-agent scenarios, algorithmic techniques to solve problems with preferences and constraints and stable matching problems. Word Embedding, Topic Modeling, Time Dependent Meaning.

74. Guidelines on using AI-powered chatbots in education and research

Guidelines on using AI-powered chatbots in education and research

The guidelines presented below are intended to offer guidance to teachers, researchers and various bodies within the university about how to relate to AI chatbots. The guidelines are developed by the university's working group on AI-powered chatbots.

- [AI-chatbots and examination](#)
- [Suspicion of cheating](#)
- [Use of AI chatbots by teachers and students during courses](#)
- [Use of AI chatbots by researchers in research and for research applications](#)
- [How will SU continue to work on the issue](#)
- [Help and support on AI chatbot issues](#)
- [Useful links](#)

Artificial intelligence tools and systems are developing fast, and recently so-called AI chatbots, such as ChatGPT, have been a topic of discussion across higher education. Chatbots are based on predictive language models and can generate coherently formulated and semantically correct text. Chatbots are trained on large language and information databases and can handle several languages. In this text, other types of systems for generating AI-based text are denoted as AI chatbots. Note that the term text is used below in a broad sense, and can refer to ordinary texts, but also code and the like.

The launch of ChatGPT has greatly increased accessibility to AI-generated text and prompts universities to ask questions about how higher education institutions should engage with these types of tools. They can provide opportunities for both research and education, but must be handled with careful judgment, and incorrect use can count as cheating and be misleading. Especially in education settings, there are concerns that students use AI chatbots to generate answers in connection with examinations.

AI chatbots, and their integration into other systems, are developing at a fast rate, and going forward the university must engage in continuous dialogue about their use in our operations. The guidelines presented in this document are intended to offer guidance to teachers, researchers and various bodies within the university about how to relate to AI chatbots. Operational responsibility for issues related to the use of AI chatbots is handled at departmental level, while the overall responsibility for handling the consequences for education and research rests with the university's decision-making bodies.

AI chatbots and examination

AI chatbots can potentially be used by students in conjunction with many different types of examinations. This may involve letting AI chatbots write texts, but it can also relate to using AI chatbots to improve texts, find errors, synthesize and present an overview of a subject area. Stockholm University therefore recommends that teachers/examiners need to decide which type of use is allowed, and which is considered impermissible, which may differ from course to course. Examples of areas of use that should be addressed in such a clarification are:

- having an AI chatbot write a text that is more or less unedited and submitted as the student's own in an attempt to mislead the examiner. This can be equated with ghostwriting or plagiarism and is normally considered cheating.
- let an AI chatbot improve a text, suggest improvements, find errors in texts, or synthesize and presents an overview of an area. In the event that one or some of these uses are permitted, students should be required to explain how the AI chatbot has been used in the production of the text.
- student use AI chatbots for peer review/opposition. If this is allowed, it should be made clear that the student must make their own assessment, and not use the results of the AI chatbot in an unprocessed manner. Even in this case, the students should explain how the AI chatbot has been used.

Reviewing forms for examination to avoid inappropriate use of AI chatbots. Examples of measures that can be taken are:

- avoid unsupervised home examination. If used, such examinations can be supplemented with another additional and complementary examination, e.g. oral examination or sit-down examinations.

- examine through supervised "open book" examination.
- introduce several incremental submissions in cases involving long text production/project work where the students report in different steps how the texts are developed
- using context-based and specific data that ties into course-specific or local conditions that make it more difficult to use AI chatbots
- have clear requirements that course literature and other literature (lectures) must always be referred to and, when possible, with specific references to pages.

Note that AI chatbots are evolving rapidly. Having context-based data or introducing requirements for references can therefore be measures that are not as effective in a long time perspective.

Keep in mind that a change in the examination forms may require that the syllabus needs to be revised, which takes time.

Suspicion of cheating

If it is suspected that a student has used an AI chatbot in an unauthorized way in connection with an examination, the offense must be investigated, and, if there is a well-founded suspicion of cheating, reported in the same way as in other cases of cheating. See [Guidelines for Disciplinary Matters at Stockholm University](#).

Tools that can assess the probability that a text is AI-generated have recently been developed and made public. One can assume, however, that countermeasures against this will be developed, and such already exist to some extent. These systems are therefore unlikely to be an effective countermeasure against unauthorized use of AI chatbots. Tools for detecting AI written text can possibly be used as part of an investigation in the event of a suspicion of cheating, but will need to be supplemented with other material to be able to form a basis in for example disciplinary committee cases. It can also be pointed out that the systems currently used for plagiarism control, in SU's case Ouriginal, cannot detect AI-generated text.

Use of AI chatbots by teachers and students during courses

AI chatbots are here to stay and are something both teachers and students need to relate to. It is therefore desirable as a teacher to think through how the use of AI chatbots can eventually be included in teaching. Examples could be that together with colleagues and students:

- analyze and reflect on benefits and problems with AI chatbots and the texts they generate
- critically review responses from AI chatbots and make students aware of the risk of inaccuracy and bias
- reflect on bias and how different perspectives are expressed in the automatic responses
- compare the AI chatbot's responses with those written by experts
- reflect on how different forms of knowledge are expressed and how these are valued when machines can now write text.

AI chatbots can also be helpful for teachers in, for example, planning teaching and producing teaching materials. Here it is important to bear in mind that the companies behind the chatbots do not report exactly which databases and which material the chatbots are trained on. Usually, the databases are very large, but time-limited text corpus, which can mean that the chatbots are not always updated with current information. In addition to this, problems can exist with bias and inaccuracies in the model itself. It is therefore important to emphasize that AI-generated material needs to be carefully reviewed so that it truly captures the course content and learning objectives, especially in examinations. An important aspect to consider is that submitted texts can be used for other purposes than intended. Sensitive information should never be sent to a chatbot. For the time being, the use of AI chatbots when assessing examinations is also advised against.

Use of AI chatbots by researchers in research and for research applications

Publishers and research funding bodies today have varying rules for how AI chatbots are permitted. Unauthorized use of them could, in certain contexts, possibly be classified as research misconduct. For the time being, the university therefore urges caution when using AI chatbots when writing research articles and research applications. If they are still used, it is important to be transparent about what was used and how. Check with the publishers or research funders what applies to the use of AI chatbots. Just as in a teaching context, it is also important to never share sensitive research information with an AI tool.

How will SU continue to work on the issue

The university is following the development of AI chatbots and other text-generating tools closely, and collaborates with other stakeholders, for example SUHF and UKÄ. The university will also

review, for example, regulations for education and examination as well as the need for skills development in the area.

Help and support on AI chatbot issues

The Center for the Advancement of University Teaching (CeUL) can provide support and assistance on educational issues about AI chatbots. CeUL Torget in Athena is also an arena where the issue can be discussed with other teachers.

In the case of questions of a more area-specific nature, for example about regulations etcetera, the Advisory committees for undergraduate studies in each Scientific area can be contacted via the faculty offices.

75. DTU opens up for the use of artificial intelligence in teaching

DTU is opening up for teachers to allow digital tools and artificial intelligence, AI, such as ChatGPT, in teaching and, in the longer term, in exams. AI is currently used in several courses and in research at DTU, and by 2024, the new technologies can be expanded and used more systematically. The decision is part of DTU's mission to develop and utilize natural and technical science for the benefit of society and to provide the best engineering education in Europe.

The more systematic use of AI challenges teachers, who have to reorganize questions in assignments and exams, but also students, who now have to assess whether AI is a relevant tool and provide information about citations and use of the technology in exams.

"DTU's task is to prepare students and ourselves to make use of new technology. We live in a digitalized world where AI is becoming increasingly widespread and accessible. Artificial intelligence makes the impossible possible in many areas, so of course it should be part of the future engineers' toolbox. This is also the attitude I see in the teachers at DTU. They find solutions and utilize technology wherever it can help strengthen students' learning and competencies," says Lars D. Christoffersen, Dean of DTU.

Students welcome AI

DTU has prepared a guide for teachers and students on the use of AI. DTU's student organization Polyteknisk Forening has been involved in the work on the guide, and the association's chairperson Natasha Hougaard is excited about the announcement on AI.

"We think it's great progress that DTU will allow AI. It shows that the university is embracing the new technology instead of being afraid of it. As students, we are prepared for the fact that AI will accelerate development where we will be met with higher academic demands because we will have access to more assistive technology. But we agree with DTU's guide and trust that the teachers know which courses and exams it makes sense to allow AI, in the same way as they consider other assistive technology," says Natasha Hougaard.

Guide for AI

DTU's guide for AI are based on a fundamental trust in the students and that they take full responsibility for the work they hand in or submit for an exam. DTU students sign DTU's honor code, which states that students must always be able to vouch for their own work and that they must never copy the work of others without acknowledging the source. Assignments may contain work done by others or produced with AI, but students must mark it clearly and the quotes they use must be correct.

Teaching needs to be personalized

The decision to integrate AI as an engineering tool will lead to significant changes in the organization of teaching at DTU. Teachers will need to review whether they need to adjust their teaching and criteria for assessing students' assignments.

In the longer term, exam questions must be set in a way that takes into account that AI can solve the factual information and utilize all available knowledge on chatbots and search engines on the internet. The wording of questions in individual tasks must therefore be formulated in such a way that problems cannot be answered using AI alone.

The development will also create new opportunities for teachers to make teaching more individual by using digitalization and AI in their pedagogical approach.

The adaptation of exam forms and questions is expected to require a transition period of six months to a year and will therefore not be an option to open up to AI across the entire academic breadth for exams until the end of 2024.

76. AI chatbots in unsupervised assessment

TU Delft Assessment Taskforce has created practical guidelines on how lecturers can deal with the influence of AI chatbots on unsupervised assessments and what mitigating measures can be taken. In June 2023, the Special Interest Group AI & Assessment updated the guidelines.

About AI tools

AI tools like the chatbot ‘chatGPT’ can produce convincing answers to certain questions, depending on their nature. However, their output is not always reliable: outputs can contain convincingly presented factual errors (so-called hallucinations). Furthermore, their training data can be outdated. For example, ChatGPT currently uses training data up to September 2021, whereas others (such as Bing) do have access to recent data. It is also important to note that most chatbots do not list their resources (Bing Chat does, though).

On the positive side, chatbots can help with:

1. Checking grammar, spelling, and references in a text
2. Generating ideas by listing information from different sources in an accessible way
3. Giving feedback
4. Summarising or expanding texts and concepts
5. Coding in a wide variety of computer languages

Use by lecturers for assessments: AI chatbots can help lecturers in creating assessments (including different versions of an assignment), answer models, and rubrics.

Guiding principles

The following assumptions have been used as the basis of the practical guidelines in the how-to section:

1. Students' use

We are assuming that AI tools are used by our students and graduates, especially because these services are currently free of charge.

2. Academic integrity

We assume that students and employees act according to principles of academic integrity, as formulated in the Code of Conduct. To foster common understanding and clarify expectations, discussions with students about integrity in the context of AI tools use are recommended.

3. Quality requirements for assessment

This is how the [quality assessment requirements](#) might be impacted by AI tools:

1. Validity & reliability:

1. The ability to use AI tools may influence the validity of the grade, because the grade may be a poorer representation of how students master the learning objectives.
 2. Allowing the use of AI tools while changing the assessment criteria (or their weight) but not (yet) the learning objectives may diminish the validity.
 3. In case students use AI tools while the teaching staff does not anticipate this, the use of AI tools may increase the grades of students. On the other hand, uncertainty about the use of AI tools may lead examiners to compensate in their assessment for (unfounded) suspicions about the use of AI tools, making the assessment inconsistent and therefore less reliable.
2. **Feasibility for students:** Mitigating measures could increase the number of assessments and therefore increase the study load.
 3. **Feasibility for teaching staff:** Extra assessments (see point c) will also increase the workload for teaching staff.
 4. **Transparency:** In addition, teaching staff may forget to communicate some of their changed expectations to students.

4. AI chatbot detection

It is currently unknown what the reliability and validity of AI chatbot detectors is.

5. Definition of fraud

The definition of fraud is (source: [model R&G](#), article 7.1):

“Fraud is taken to mean any act or omission by a student that makes it fully or partially impossible to properly assess the knowledge, insight and skill of that student or another student.

6.

Attribution

The use of AI chatbots (and of tools in general) should be acknowledged and properly referenced, to ensure the distinction between the students' original ideas and those provided by AI, and to check whether the student critically checked the output of the AI-generated outcome. However, this challenging, as it is expected that AI tools are to be used in an organic and evolutionary manner.

7.

Accessibility

Currently, most chatbots are still free-of-charge, which makes a low threshold for students to use this. In the (near) future, it is likely that users will need to pay a fee. This could potentially lead to the need for higher education institutions to accommodate these AI tools when they are actively used in our education and that all students have equal access to these types of tooling.

8.

Security & privacy

As AI tools use user input to train future versions, it can have consequences for the privacy and intellectual property of information that is fed to the AI tool during its use. Almost every online tool requires the use of personal data.

ChatGPT specific information: OpenAI is the company that offers the ChatGPT service. Users need to sign up for an account in order to use <https://chat.openai.com>. Besides user account information OpenAI also processes the following personal information:

1. User Content: when you use ChatGPT, OpenAI will collect personal Information that is included in the input, file uploads, or feedback that you provide to ChatGPT during interactions
2. IP-Addresses
3. Browser user agents
4. Operating System and device information
5. Cookies
6. Tracking Identifiers

OpenAI uses technology requires sharing personal data with third parties and that data is stored on servers in the USA. At this moment is unclear with whom personal data is shared and what parties are responsible for the use and protection of your personal data. The security and privacy team will update when new information becomes available.

8. Ethical issues

In addition, ethical questions are arising regarding the current and future influence of AI chatbots on truth finding and society as a whole, as well as regarding the power of its owners (big tech companies) and the impact of the technology on vulnerable communities (exploited labour) and the environment.

9. Rapid evaluation of AI tools

Many of the current shortcomings will be (partially) solved in the next versions of AI tools. Therefore, it is important to focus on the possibilities and not so much on current shortcomings, because the latter change. However, we should consider the more static risks of these technologies, which are unlikely to change. In other words, we should distinguish between shortcomings and risks.

How to assess assignments and projects

Invigilated exams versus assignments and projects

During classical written exams and digital exams in the [TUD secure exam environment](#), students do not have access to the internet, and therefore your students cannot access online AI tools. The same holds for oral exams that are held on campus.

On the other hand, if students work on assignments (exam-like or other) outside an exam hall and without invigilators (Dutch: surveillanten), the use of AI tools cannot be prevented.

Advice for fraud prevention in (non-invigilated) assessment

1. Discover possibilities and limitations and discuss them with your students

Feed your assignments to chatbots and study their output. How would you assess the output using your answer model or rubric? You can use this information to get a feeling of whether students used AI tools in their work.

Discuss the possibilities and limitations with your students of using AI tools in unsupervised assignments. Let students use the AI tool and let them reflect on the answer provided. Train students to not trust the answer of AI tools, even for questions that are not too difficult and require mostly factual knowledge. Students need to internalize that they need to double-check all output of AI tools, to prevent them from learning incorrect facts and reasoning.

2. Safe use of AI tools and plugins

TU Delft recognises the value of AI tools but sharing data is never without any risks. If you choose to use AI tools like ChatGPT or AI plugins, we recommend you take the following recommendations to heart:

Reveal nothing: Do not share any personal data, internal information or (highly) confidential information during your interactions with the AI tool.

1. **Reveal nothing:** Do not share any personal data, internal information or (highly) confidential information during your interactions with the AI tool.
2. **Private/incognito window:** Use ChatGPT while browsing in a private or incognito window.
3. **Password:** As long as AI tools do not offer Single Sign On, account and password management is up to the individual user. Do not reuse passwords, preferably use a password manager or other ways to create a safe password and change the password on a regular basis.
4. **AI plugin awareness:** The new plugin functionality of, for example, ChatGPT offers the possibility to include external sources. This makes it easier to share data. Keep in mind that the above recommendations also apply to these plugins.

3. Be transparent and explain your choices

If you consider the use of AI bots in your course detrimental to achieving the learning objectives, clearly state your reasons to your students. Make sure they understand that they might fail in the summative assignment when they do not have access to these AI tools. Refer students back to the definition of fraud and to our [TU Delft code of conduct](#).

4. Attribute correctly

Inform your students on how you expect them to correctly attribute the use of AI tools.
Examples:

1. **Reflection:** Have them write a short section on how they used chatbots and in what ways it was and was not helpful, and what they learned.
2. **Coding:** Give instructions on using AI tools for developing software code, and on how to acknowledge their use.

5. Reduce the need for students to rely on AI tools by making them feel confident

1. **Have sufficient feedback moments** during the course and ask students to reflect on how they processed the feedback. If possible, do this in a discussion.
2. **Regularly check the progress of individual students** during their projects/assignments (if feasible). This is also good for learning and confidence building if you turn it into a supervision/feedback moment. Check during the creation of the deliverable (of a project/assignment) whether they are all contributing / learning, for example by brief oral discussions of a product they are working on (e.g. after finishing a specific (small) step in a project/code assignment).

6. Focus on the process

1. **Shift assessment criteria towards the process** instead of the deliverable. Make sure that the assessment is valid and transparent for students.
2. **Version control:** Track the progress of students through version control in e.g. Word or Gitlab. Are they processing feedback proactively?

7. Take fraud detection measures

Take fraud detection measures and report suspicions of fraud to your Board of Examiners:

1. **Oral authenticity check:** Do an [oral authenticity check](#) (4a) to check if it is likely that the student produced the text in themselves. This should either be a random check, or based on justifiable parameters to prevent bias.
2. **Check the transfer of skills & knowledge:** Consider adding a written exam to a project in which students have to answer questions on a case that is similar to their project. That way, you can test the students' ability to transfer their knowledge to another situation. Additionally, this aids with retention of knowledge & skills, especially if you discuss the exam in class. Carefully consider the timing and the weight of the exam (consider making the exam pass-fail) to prevent students from focussing on the exam instead of on their project. Adding assessments adds to the study load and is not permitted during the academic year without the permission of the board of examiners (and the programme director).

8. Long run – rethink your course

Rethink your course's assessment plan. If necessary, adjust the learning objectives. This doesn't necessarily mean that the taxonomy level should be increased since this could lead to an increase

in the difficulty and study load of the course. Consider the relation to other courses in your decision.

Keep in mind that these changes require a study-guide adjustment and will therefore have to be approved by your programme director, the Board of Studies and the Faculty Student Council before the start of the academic year. Changes during the academic year can only occur in very special circumstances after approval by the board of examiners, see [here](#)).

77. Artificial Intelligence in education

Técnico presents resolution on the use of tools such as ChatGPT

The School does not prohibit the use of these tools and encourages students and professors to use them as learning and teaching assistants.

“It’s the most exciting topic of recent years”, said Rogério Colaço. This is how the president of Instituto Superior Técnico referred to the use of artificial intelligence (AI) tools, in his opening speech, at the 6th session of Contigo+ Programme.

The event took place on January 9 at the Técnico Congress Centre, Alameda campus, during which the [Pedagogical Council presented the Técnico’s deliberation on the use of AI language models](#) (such as ChatGPT) by students and professors. According to the document drawn up by the Pedagogical Council, “students should be encouraged to use these tools as a learning and work assistant”. The text also states that “professors should use AI-based tools to enrich, simplify [and] update the teaching process” and “no general prohibition should be adopted regarding the use of AI tools in teaching or assessment methods”.

Later in his speech, Rogério Colaço stressed: “we’re not in the 80s or 90s of the last century – today’s students are different from those we had 40, 30, 20, five or six years ago”. According to the president of Técnico, the most striking difference is that these students “have grown up with different access to information”. He also recalled that unlike a hierarchical model of teaching

where the teacher holds all the knowledge, today “students have the same information, or more because they have a better grasp of technology”.

Teresa Peña, president of the Pedagogical Council, emphasised the need to combine these tools “with traditional assessment methods with increasing interaction between professor and student – more oral exams, more presentations, more discussions”. She also called for the use (or not) of AI tools to be made explicit in the assessment method. She also talked about the limited reliability of these tools, reminding users of the need to be critical.

[Arlindo Oliveira](#), professor and former president of Técnico, gave a talk on “the impact of artificial intelligence technologies on the teaching process and knowledge assessment”. The speaker emphasised the school’s role in dealing with AI developments, “Técnico has to be prepared not only for the current state of technology, but also for what will be the future of technology in five or ten years”. Highlighting the power of AI in education, he pointed out that “the capacity of these systems as personalised tutors should not be underestimated and we must be able to take the most of it”.

At the end of the talk, there was an exchange of ideas between the speaker and the audience, focused on exploring the logistical impact of using AI tools in assessment. The event also included talks by professors [Carlos Silva](#) (“Elementary applications of generative AI in teaching”) and [João Ferreira](#) (“Using AI to Increase Productivity in Teaching”) and respective exchange of ideas.

The Pedagogical Council’s deliberation follows the work carried out by the REFLeT Commission (‘Reflection on Teaching and Training in the Era of Large Language Models’), which included a group of experts who produced a [report on this topic](#) in July. Chaired by Arlindo Oliveira, the commission brought together members of the Scientific Council ([Eduardo Júlio](#); [Pedro Lima](#)) and the Pedagogical Council (Carlos Silva; [João Pimentel Nunes](#)), all of them Técnico professors, as well as people from outside Técnico (Paulo Mota Pinto and Porfírio Silva).

78. Guidelines for the use of generative AI

Guidelines for the use of generative AI

AI (Artificial Intelligence) refers to a machine's ability to exhibit skills - such as reasoning, learning, planning, and creativity - that are inherent to humans. Within this domain, generative AI shines as the master of creation. It empowers machines to generate new content based on existing data, ranging from images to text and sound.

On this page, you will find the current guidelines for the use of generative AI at Utrecht University (UU). These guidelines enable us to harness the opportunities presented by generative AI while simultaneously mitigating the risks associated with its use.

What can generative AI be used for?

Generative AI can be applied in many different areas, making it a potentially valuable tool for various kinds of use.

- Written communication
- Optimisation and automation
- Creative applications
- Data visualisation
- Software development
- Search Engine Optimisation (SEO)

AI for science communication

AI language models like ChatGPT can be utilized as tools in science communication because they can explain complex scientific concepts in an accessible and understandable manner.

[CLICK HERE TO READ ABOUT AI FOR SCIENCE COMMUNICATION](#)

Risks

To safely harness the possibilities of generative AI, it is essential to recognize the risks associated with its use and proactively take measures to reduce them. Below, you will find a list of the main risks associated with the use of Generative AI.

Privacy

The main privacy concerns related to AI involve the risk of data breaches and unauthorized access to personal information. Given the volume of data collected and processed, there is a risk that it could end up in the wrong hands, for instance, through hacking.

Copyright

Generative AI can generate content that infringes on copyright by creating material based on existing copyrighted works without the permission of the rights holders. This can lead to legal issues and claims of infringement. If AI generates content closely resembling existing works, it may result in accusations of plagiarism, where the original copyright holders may contend that their work has been unlawfully copied.

Lack of responsibility and accountability

Because generative AI can autonomously generate content, it is sometimes challenging to determine who is responsible for the generated content. Various parties, including the developer of the AI tool, the user, or even the AI itself, may be involved.

Incorrect or discriminatory information

Generative AI relies on training data and may contain biases that can manifest in the generated output. This can lead to discrimination or inequality, for example, in the case of job applications or decision-making. Additionally, the technology can generate output of lower quality or incorrect results.

Dependency on technology

Dependency on generative AI can result in employees losing certain skills and expertise. Additionally, the use of AI may lead to job loss in areas that can be automated. It is important to find a balance between the use of generative AI and the preservation of human knowledge and insights.

Safety

Generative AI can be exploited by malicious actors, for instance, to forge documents or manipulate image and audio content (deepfakes). This can pose security risks, such as identity fraud.

Guidelines for the use of generative AI

When using generative AI applications as an employee of Utrecht University, you are required to adhere to the following guidelines.

a. Privacy en data protection

Do not share sensitive personal and/or company data when using generative AI applications.

Utrecht University handles personal data carefully and operates within the boundaries of the law, specifically the General Data Protection Regulation (GDPR).

Personal data is not rented, sold, or otherwise shared with or disclosed to third parties.

UTRECHT UNIVERSITY PRIVACY STATEMENT

EXTERNAL LINK

b. Intellectual property

Every AI-generated expression must be thoroughly checked to ensure that the content complies with applicable copyright laws and that any required permissions or licenses have been obtained correctly.

Online plagiarism detection tools are available to help identify potential duplicates of existing content. Popular plagiarism detection tools include Turnitin, Copyscape, and Grammarly.

When using specific information, quotes, or ideas from a copyrighted work, the source and author must be acknowledged.

READ MORE ABOUT COPYRIGHT

EXTERNAL LINK

CHECKING IMAGES FOR COPYRIGHT

EXTERNAL LINK

c. Accountability and Transparency

All AI-generated output must be thoroughly checked to prevent the dissemination of incorrect information and/or discriminatory content.

Under no circumstances should generative AI be used for illegal, harmful, and/or discriminatory activities.

When using generative AI in your duties as an employee of Utrecht University, you are deemed personally responsible for the generated content. This means that you are liable in case of damage or unwanted consequences resulting from the use of AI.

It is crucial for employees to acknowledge the use of AI in their duties. If any form of AI is used, the tools employed must be disclosed.

79. Guidelines for the Use of AI Tools

Guidelines for the Use of AI Tools Disclaimer: These guidelines have been adapted from a policy statement issued by the University of Portsmouth. (Points 1, 3, and 5 have been taken over verbatim.) Artificial Intelligence (AI) tools such as ChatGPT use algorithms to generate output (such as text, images and code), based on user questions or commands. The seven principles listed below will help you determine under which circumstances the use of AI tools is appropriate – and when it is not. Adhering to these principles will enable you to make use of AI tools in a way that complements and enhances your studies, without risking academic integrity. Seven Principles for the Use of AI Tools

1. Use AI as a tool to assist and inform you in your initial research, generation of ideas, planning and output development, but not as a replacement for your critical thinking and analysis.
2. Ensure that you appropriately cite and reference any text or output generated by AI in your assignment, along with any other sources you use (see writer's manuals for Literature, Linguistics, and Language Skills and Culture for details). You must indicate clearly where in your assessment task you have used AI-generated material.
3. Understand the AI tool's limitations and therefore use it in conjunction with other sources to ensure the information you present is credible and reliable. You need to check the accuracy of all information generated by AI tools.
4. Be aware of the UZH's and the English Department's regulations, such as the Code of Honor for online exams (<https://www.zi.uzh.ch/en/support/e-learning-and-assessmentsupport/online-exams/honor-code.html>) or the relevant guidelines concerning plagiarism (Rahmenverordnung über die Bachelor- und Masterstudiengänge an der Philosophischen Fakultät der Universität Zürich, RVO PhF, 27.08.2018, §11 and §12; see also the PDF document “What Constitutes Plagiarism?”: <https://www.es.uzh.ch/dam/jcr:20ac4c8d-35b8-4a39-8b36-fb471ff330a1/What%20Constitutes%20Plagiarism.pdf>).
5. Make sure that any final product (your assessment as submitted) is your own work, and not just copied from an AI generator, in whole or in part. You can use the generated text or output as a starting point to give you inspiration or guidance, but the final submitted assessment must be all your work, your creation, and your analysis.

6. Remember that writing is not just a way of putting your fully formed, finished thoughts on paper. Rather, writing is a way of testing, reshaping, and refining your ideas. While it may make sense to use AI tools to help you solve specific problems, you should never use them to ‘circumvent’ the challenge of expressing your ideas in your own words.

7. Bear in mind that any kind of work you do comes with ethical obligations on your part: AI tools can respond to queries, but it is you who will be responsible for any material you decide to use – including, for example, biased findings or other types of misinformation.

80. Russell Group principles on the use of generative AI tools in education

Russell Group principles on the use of generative AI tools in education Our universities are committed to the ethical and responsible use of generative AI and to preparing our staff and students to be leaders in an increasingly AI-enabled world. The rise of generative artificial intelligence (AI) has the potential for a profound impact on the ways in which we teach, learn, assess, and access education. Our universities wish to ensure that generative AI tools can be used for the benefit of students and staff – enhancing teaching practices and student learning experiences, ensuring students develop skills for the future within an ethical framework, and enabling educators to benefit from efficiencies to develop innovative methods of teaching. Valuable work undertaken by organisations such as the Quality Assurance Agency for Higher Education (QAA) and Jisc has helped develop the sector’s understanding of the opportunities and considerations of generative AI¹², and the Department for Education (DfE) has set out its position on the use of generative AI in the pre-university education sector³. Russell Group universities have contributed sector-wide insight and have been proactively working with experts to revise and develop policies that provide guidance to students and staff. Collaboration, coordination, and consistency on this issue across the education and professional sectors – including professional bodies, schools, FE colleges and employers – will be crucial. In recognition of this, Russell Group universities have collectively developed the following principles that will guide the approach to generative AI tools across our universities and, we hope, beyond:

- Universities will support students and staff to become AI-literate.
- Staff should be equipped to support students to use generative AI tools effectively and appropriately in their learning experience.
- Universities will adapt teaching and assessment to incorporate the ethical use of generative AI and support equal

access. 4. Universities will ensure academic rigour and integrity is upheld. 5. Universities will work collaboratively to share best practice as the technology and its application in education evolves.

1. Universities will support students and staff to become AI-literate.

1.1 Generative AI tools are capable of processing vast amounts of information to generate responses but they have significant limitations. It is important that all students and staff understand the opportunities, limitations and ethical issues associated with the use of these tools and can apply what they have learned as the capabilities of generative AI develop. These include: (a) Privacy and data considerations: whether a generative AI tool is designed to learn directly from its users' inputs or not, there are risks to privacy and intellectual property associated with the information that students and staff may enter. (b) Potential for bias: generative AI tools produce answers based on information generated by humans which may contain societal biases and stereotypes which, in-turn, may be replicated in the generative AI tool's response. (c) Inaccuracy and misinterpretation of information: data and information contained within generative AI tools is garnered from a wide range of sources, including those that are poorly referenced or incorrect. Similarly, unclear commands or information may be misinterpreted by generative AI tools and produce incorrect, irrelevant or out-of-date information. This means that accountability for the accuracy of information generated by these tools when transferred to another context lies with the user. (d) Ethics codes: users of generative AI tools should be aware that while ethics codes exist, they may not be embedded within all generative AI tools and that their incorporation, or otherwise, may not be something that users can easily verify. (e) Plagiarism: generative AI tools re-present information developed by others and so there is the risk of plagiarised content and/or copyright infringement being submitted by a user as their own, and artwork used by image generators may have been included without the creator's consent or licence. (f) Exploitation: the process by which generative AI tools are built can present ethical issues. For example, some developers have outsourced data labelling to low-wage workers in poor conditions⁴. 1.2 Our universities will provide guidance and training to help students and staff understand how generative AI tools work, where they can add value and personalise learning, as well as their limitations. By increasing AI-literacy, our universities will equip students with the skills needed to use these tools appropriately throughout their studies and future careers, and ensure staff have the necessary skills and knowledge to deploy these tools to support student learning and adapt teaching pedagogies.

2. Staff should be equipped to support students to use generative AI tools effectively and appropriately in their learning experience.

2.1 Our universities will develop resources and training opportunities, so that staff are able to provide students with clear guidance on how to use generative AI to support their learning, assignments, and research. 2.2 The appropriate uses of generative AI tools are likely to differ between academic disciplines and will be informed by policies and guidance from subject associations, therefore universities will encourage academic departments to apply institution-wide policies within their own context. Universities will also be encouraged to consider how these tools might be applied appropriately for different student groups or those with specific learning needs. 2.3 Engagement and dialogue between academic staff and students will be important to establish a shared understanding of the appropriate use of generative AI tools. Ensuring this dialogue is regular and ongoing will be vital given the pace at which generative AI is evolving.

3. Universities will adapt teaching and assessment to incorporate the ethical use of generative AI and support equal access.

3.1 Universities continually update and enhance their pedagogies and assessment methods in response to drivers including new research, technological developments and workforce needs – adapting to the use of generative AI technology is no different. Incorporating the use of generative AI tools into teaching methods and assessments has the potential to enhance the student learning experience, improve critical reasoning skills and prepare students for the real-world applications of the generative AI technologies they will encounter beyond university. 3.2 Appropriate adaptations to teaching and assessment methods will vary by university and discipline, and protecting this autonomy is vital. All staff who support student learning should be empowered to design teaching sessions, materials and assessments that incorporate the creative use of generative AI tools where appropriate. Professional bodies will also have an important role in supporting universities to adapt their practices, particularly in relation to accreditation. 3.3 As the technologies develop and new generative tools become available, elements of generative AI used within universities may reside behind paywalls or be restricted to paying subscribers. Universities will need to consider how best to respond to a potential proliferation of such subscription tools and attempt to ensure fairness of access so that students and staff can access the generative AI tools and computing resources they need in support of their teaching and learning practices.

4. Universities will ensure academic rigour and integrity is upheld.

4.1 All 24 Russell Group universities have reviewed their academic conduct policies and guidance to reflect the emergence of generative AI. These policies make it clear to students and staff where the use of generative AI is inappropriate, and are intended to support them in making informed decisions and to empower them to use these tools appropriately and acknowledge their use where necessary. 4.2 Such clear and transparent policies are critical to maintaining consistent and high standards of learning, teaching and assessment across Russell Group universities. 4.3 Ensuring academic integrity and the ethical use of generative AI can also be achieved by cultivating an environment where students can ask questions about specific cases of their use and discuss the associated challenges openly and without fear of penalisation.

5. Universities will work collaboratively to share best practice as the technology and its application in education evolves.

5.1 Navigating this ever-changing landscape will require collaboration between universities, students, schools, FE colleges, employers, sector and professional bodies, with the ongoing review and evaluation of policies, principles and their practical implementation. 5.2 Our universities will regularly evaluate policies and guidance for staff and students relating to generative AI tools and their impact on teaching, learning, and assessment practices. This will include monitoring the effectiveness, fairness, and ethical implications of the integration of generative AI tools into academic life, and adapting policies and procedures to ensure they remain valid as generative AI technologies evolve. 5.3 Fostering relationships between higher education institutions, schools, employers, professional bodies who accredit degrees, AI experts, leading academics and researchers, as well as ensuring an inter-disciplinary approach to addressing emerging challenges and promoting the ethical use of generative AI, will be crucial. Russell Group universities recognise the challenges that lie ahead and will continue to value the input of others, along with contributing expertise to the national and international discussions around generative AI and its applications within teaching, learning, assessment and support.

Funding: This research is funded by the National Science Foundation under grant number 2125858.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

